A 0610-1

Public Opinion Quarterly

Published for the American Association for Public Opinion Research by the University of Chicago Press

J.-N. KAPFERER	A Mass Poisoning Rumor in Europe
RUTH THIBODEAU	Blacks in *New Yorker* Cartoons
ROGER TOURANGEAU and others	Carryover Effects in Attitude Surveys
LAWRENCE F. FEICK	Latent Class Analysis of Questions with Don't Know Responses
CAROL S. ANESHENSEL and others	Panel Survey of Mexican American Adolescents
THE POLLS	Review: The 1988 Presidential Primaries *by Larry M. Bartels and C. Anthony Broh* Report: 1960–1988 New Hampshire Presidential Primaries *by Kenneth E. John* Response: Flaw in the Review and the Review Process *by Jon D. Miller*
BOOKS	Reviews *by Robert M. Groves, Stanley Kelley, Jr., Barbara A. Bailar, and Warren J. Mitofsky*

Winter 1989

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Leo Bogart, Chair (1985–89)
Herbert Abelson (1988–92) Eleanor Singer (1987–91)
David Sills (1986–90)

Gladys Engel Lang and Joan S. Black,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor David Carlson
Office Manager Carlotta Nosel
Assistant to the Editor Shoshana Colman

The Polls
Reviews: Stanley Presser, Seymour Sudman
Reports: Tom W. Smith

EDITORIAL BOARD

Robert P. Abelson
Barbara A. Bailar
Sandra J. Ball-Rokeach
George F. Bishop
Lawrence D. Bobo
Thomas D. Cook
Irving Crespi
James A. Davis
Stanley Feldman
Norval D. Glenn
Robert M. Groves
Deborah R. Hensler
Shanto Iyengar
Jon A. Krosnick
Kurt Lang
Michael MacKuen
McKee J. McClendon
Jack M. McLeod
John Modell
John Mueller
Harold Nisselson
Steven J. Rosenstone
Norbert Schwarz
David O. Sears
Robert Y. Shapiro

Public Opinion Quarterly (ISSN 0033-362X) Volume #53, 4 issues (1989)

© 1989 by the American Association for Public Opinion Research. Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

465

PUBLIC OPINION QUARTERLY

VOLUME 53 / WINTER 1989 / NUMBER 4

Articles

A Mass Poisoning Rumor in Europe 467
J. N. KAPFERER

From Racism to Tokenism: The Changing Face of Blacks in *New Yorker* Cartoons 482
RUTH THIBODEAU

Carryover Effects in Attitude Surveys 495
ROGER TOURANGEAU, KENNETH A. RASINSKI, NORMAN BRADBURN, and ROY D'ANDRADE

Latent Class Analysis of Survey Questions That Include Don't Know Responses 525
LAWRENCE F. FEICK

Participation of Mexican American Female Adolescents in a Longitudinal Panel Survey 548
CAROL S. ANESHENSEL, ROSINA M. BECERRA, EVE P. FIELDER, and ROBERLEIGH H. SCHULER

The Polls

Review: The 1988 Presidential Primaries 563
LARRY M. BARTELS and C. ANTHONY BROH

Report: 1980–1988 New Hampshire Presidential Primary Polls 590
KENNETH E. JOHN

Response: The Flaw Is in the Review and the Review Process 606
JON D. MILLER

Books

John Goyder, *The Silent Minority* 610
ROBERT M. GROVES

466

Irving Crespi, *Pre-Election Polling: Sources of Accuracy and Error* 613
STANLEY KELLEY, JR.

Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, and Joseph Waksberg, *Telephone Survey Methodology* 615
BARBARA A. BAILAR

Herbert Asher, *Polling and the Public: What Every Citizen Should Know* 617
WARREN J. MITOFSKY

Book Notes 619

NOTICE TO CONTRIBUTORS 621

INDEX TO VOLUME 53, 1989 623

A MASS POISONING RUMOR IN EUROPE

J. N. KAPFERER

Abstract For more than ten years, a rumor has been spreading in Europe. Transmitted from hand to hand on a leaflet, this rumor accuses ten well-known brands of food products of being toxic and producing cancer. Because of its persistence, extensiveness, and tangible nature, the rumor lends itself to empirical research. This article presents the most significant results of a number of studies assessing the rumor's penetration, its modes of diffusion, and its behavioral effects.

For more than ten years, a contamination rumor has been circulating in France. Long-lived, despite all official denials, this rumor is passed along not by word of mouth but from hand to hand, in the form of a leaflet. Thus, by its tangibility and durability, this rumor lends itself to empirical investigations. Such analyses are all the more necessary in so far as this leaflet has gained considerable terrain: in France, it is estimated that about half of all housewives have been exposed to this rumor. In other words, by its persistence, pervasiveness, and persuasiveness, this rumor—known as "the leaflet of Villejuif"—is one of the most important rumors France has had for many years. It has also spread to many other European countries; translated versions are now circulating in Great Britain, Germany, Italy, and abroad in the Middle East and Africa.

A sign of a rampant mistrust of modern food and technology, this rumor testifies to the vulnerability of public opinion across all social classes. This article is divided into two parts: the first part presents the facts and the second draws implications from such an extensive phenomenon.

History and Diffusion of the Leaflet of Villejuif

In the spring of 1976, the French headquarters of major food companies operating in France (Coca-Cola, Cadbury-Schweppes, Martini,

J. N. KAPFERER is President of the Foundation for the Study and Information on Rumors (Paris, France) and Professor of Communication at l'Ecole des Hautes Etudes Commerciales and l'Institut Supérieur des Affaires, France.

Public Opinion Quarterly Volume 53:467–481 © 1989 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/89/0053-04/$2.50

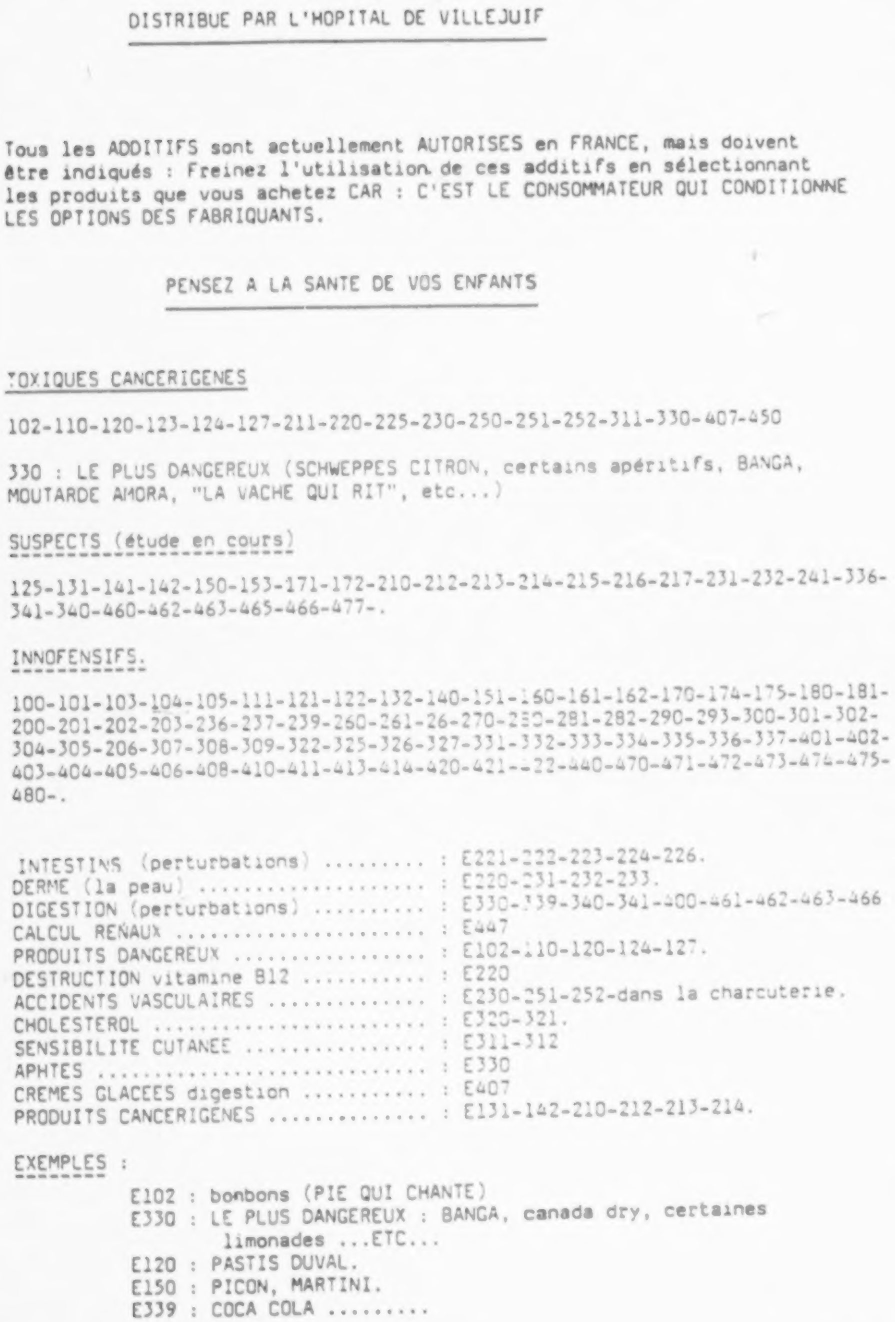
BSN, Gervais-Danone, and others) received from their local salesmen a leaflet, handed to them by a friend, colleague, or some local retailer who had just received it. Typed on an ordinary typewriter, this leaflet urged consumers to boycott ten popular brands of food or beverages: Coke, Schweppes, Martini, the Amora Mustard, Banga orange juice, etc. According to the leaflet, these brands contained additives which despite being authorized in France actually would be toxic or carcinogenic. To warrant this fear, the leaflet referred to an anonymous source: a "Hospital in Paris" (there are 37 hospitals in Paris) having conducted research on food additives. The leaflet presented a threefold classification of all food additives "authorized for consumption by the French Ministry of Health"; 17 were said to be "toxic and carcinogenic," 27 were "suspect" and under "current scrutiny," and the others were labeled "innocuous." To encourage consumers to boycott the dangerous brands and products, the leaflet lists all food additives pertaining to these three categories, each additive being referred to by its code name.[1]

Soon new versions of the leaflet appeared during the same year, spontaneously passed on from hand to hand, each of them typed on different but still ordinary typewriters. They were identical word-for-word with the first version, with two major exceptions: new brands were added to the list of those to be boycotted, and the attributed source became more precise. The mention of a "Paris hospital specializing in cancer research" was replaced by an explicit attribution to the Hospital of Villejuif. Villejuif is the name of a small city on the outskirts of Paris whose hospital is internationally famous for its advanced research on cancer (Figure 1).

Such explicit reference prompted internal inquiries at the hospital. No researcher recognized any relationship with the leaflet. The Hospital of Villejuif issued repeated official denials pointing out the usurpation of its name for fraudulent purposes. In fact, toxicologists and cancer researchers who examined the leaflet soon identified its deceptive content: many additives known to be dangerous, hence unauthorized in France, were declared innocuous by the leaflet. Conversely, the additives declared toxic by the leaflet were held to be innocuous by health officials and experts. For instance, the leaflet points out E 330 as the most dangerous and carcinogenic food additive of all. The incongruous nature of such accusation emerges given the fact that E 330 is nothing but the harmless citric acid, found in citrus fruit.

---

[1] In the Common Market, in order to overcome translation difficulties, all food additives are designated by a code name, *E* plus three digits (*E* meaning Europe). For instance, citric acid is E 330 and orthophosphate of sodium is E 339. All food brands containing additives must specify their code names among the ingredients on their packaging.



DISTRIBUE PAR L'HOPITAL DE VILLEJUIF

Tous les ADDITIFS sont actuellement AUTORISES en FRANCE, mais doivent être indiqués : Freinez l'utilisation de ces additifs en sélectionnant les produits que vous achetez CAR : C'EST LE CONSOMMATEUR QUI CONDITIONNE LES OPTIONS DES FABRIQUANTS.

PENSEZ A LA SANTE DE VOS ENFANTS

TOXIQUES CANCERIGENES

102-110-120-123-124-127-211-220-225-230-250-251-252-311-330-407-450

330 : LE PLUS DANGEREUX (SCHWEPPES CITRON, certains apéritifs, BANGA, MOUTARDE AMORA, "LA VACHE QUI RIT", etc...)

SUSPECTS (étude en cours)

125-131-141-142-150-153-171-172-210-212-213-214-215-216-217-231-232-241-336-341-340-460-462-463-465-466-477-.

INNOFENSIFS.

100-101-103-104-105-111-121-122-132-140-151-160-161-162-170-174-175-180-181-200-201-202-203-236-237-239-260-261-26-270-280-281-282-290-293-300-301-302-304-305-206-307-308-309-322-325-326-327-331-332-333-334-335-336-337-401-402-403-404-405-406-408-410-411-413-414-420-421-422-440-470-471-472-473-474-475-480-.

INTESTINS (perturbations) .... : E221-222-223-224-226.
DERME (la peau) .... : E220-231-232-233.
DIGESTION (perturbations) .... : E330-339-340-341-400-461-462-463-466
CALCUL REÑAUX .... : E447
PRODUITS DANGEREUX .... : E102-110-120-124-127.
DESTRUCTION vitamine B12 .... : E220
ACCIDENTS VASCULAIRES .... : E230-251-252-dans la charcuterie.
CHOLESTEROL .... : E320-321.
SENSIBILITE CUTANEE .... : E311-312
APHTES .... : E330
CREMES GLACEES digestion .... : E407
PRODUITS CANCERIGENES .... : E131-142-210-212-213-214.

EXEMPLES :

E102 : bonbons (PIE QUI CHANTE)
E330 : LE PLUS DANGEREUX : BANGA, canada dry, certaines limonades ...ETC...
E120 : PASTIS DUVAL.
E150 : PICON, MARTINI.
E339 : COCA COLA ....

Figure 1. The Leaflet of Villejuif.

Translation: Distributed by the Villejuif Hospital. All these food additives are authorized in France. Block the usage of these additives by selecting the products you buy. It is the consumer that bears upon manufacturers' decisions. Think of your children. Toxic carcinogenic: 102-110-120 . . . Suspicious: 125-131-141 . . . Harmless: 100-101 . . .

However, few people—except the specialists—understand the meaning of the mysterious code names of all food additives. Furthermore, the general public has not the slightest knowledge of the pros and cons of each additive, even citric acid. Since the official denials were hardly heard, the leaflet's diffusion went on unchallenged, in a country where any departure from truly natural products is certain to arouse latent fears and anxieties. It was still circulating in 1988 in France; translated it has circulated in other European countries.

EXTENT AND PATTERNS OF DIFFUSION

To measure the extent of the rumor's diffusion, exposure questions were inserted in standard omnibus polls. A national representative sample of French housewives was interviewed by Institut Français d'Opinion Publique, a member of the Gallup organization in France. The sampling methodology is the quota method: sample representativeness was based on specific quotas of age, occupation of household head, region, and type of habitation (Table 1).

After three years of spontaneous diffusion, 43% of French housewives had read the leaflet. The observed decline between 1979 and 1983 is mostly due to forgetting among those exposed to the leaflet early on. Actually, the number of forgetters is larger than this decrease since new persons had an opportunity to encounter the leaflet during this period. In the 1983 poll, among those declaring having read the

Table 1. Percentage of Respondents Who Were Exposed to the Leaflet

	April 1979	April 1980	October 1983
Have read the leaflet	43%	39%	33%
Have not read the leaflet	57	61	67
Have heard about it	8	8	7
Have not heard about it	49	53	60
N	932	990	939

QUESTIONS: Have you had the opportunity, at your home, at work, or anywhere else, to read a document or leaflet warning consumers against the ingestion of additives in food, designated by the letter *E*? (yes or no)

If not, have parents, friends, or colleagues spoken to you about such a leaflet warning consumers against the additives in food? (yes or no)

Table 2. Respondents Who Were Exposed to the Leaflet, by Age, Income, and Family Size, October 1983

	Percentage	*N*
Entire sample	33%	939
Age		
15–24	41	92
25–34	39	176
35–49	44	240
50–64	31	203
65+	14	228
Monthly family income[a]		
Unknown	22	138
Below 3,000 F.F.	20	95
3,000–4,999 F.F.	23	202
5,000–6,499 F.F.	30	111
6,500–9,499 F.F.	40	212
9,500+ F.F.	52	182
Family size		
One person	17	200
Two persons	30	260
Three persons	41	186
Four persons	40	293

QUESTION: Have you had the opportunity, at your home, at work, or anywhere else, to read a document or leaflet warning consumers against the ingestion of additives in food, designated by the letter *E*? (yes or no)

[a] One U.S. dollar = six F.F.

leaflet, 7% said they did it this same year and 11% said they read it between 1980 and 1982.

As expected, the leaflet's diffusion is uneven: it followed selective paths, penetrating those most concerned by its allegations (Degh and Vazsonyi, 1975). As shown in Table 2, there is a strong relationship between exposure to the leaflet and age, socioeconomic status, and the presence of children in the household. Since it explicitly urges mothers to boycott a number of brands popular among children, the leaflet reaches high penetration rates among mothers having children of an age to be affected by these brands.

In rumor research, there are frequent allusions to rumors being mostly persuasive among the poorly educated. In his famous study in the psychology of panic, Cantril (1966:112) found that critical ability was correlated with education. Confronted by Orson Welles's "Invasion from Mars'' broadcast, about half as many people with a college education, as compared to those with grammar school training, believed the broadcast was a real news report. For Shibutani (1966:123), education becomes a crucial variable in some rumors in that it provides a better basis for judgment. In contrast with these allusions, the leaflet reached its highest penetration rates among the wealthiest (and most educated) households. This is probably related to the fact that in France the consumerist movement drew most of its troops from the educated segment. The audience of consumerist magazines is upwardly skewed: its rate of penetration increases with the level of education. In essence, the leaflet of Villejuif provides confirmatory and allegedly scientific evidence of consumerist fears; heavily advertised national brands would be ready to incorporate any food additive or coloring in order to boost sales. Search of supportive information is a basis of selective exposure to communications; hence the success of the leaflet.

THE OPINION LEADERS AND THE RUMOR

Because the leaflet made direct reference to health, cancer, and children, it was hypothesized that the public might ask opinion leaders for advice, being supposedly able to understand the very technical content of the message. Actually some people had found the leaflet posted in the waiting room of their physician or in hospitals. Many cases of primary school teachers handing out the leaflet directly to pupils in the classroom were also reported.

In order to assess physicians' and teachers' attitudes, two semiquantitative surveys were conducted at the end of 1983. One hundred physicians were randomly drawn from the physician national directory and interviewed by telephone. One hundred primary school teachers were recruited by the quota method and interviewed on a face-to-face basis. Despite the small sample sizes, the figures are unambiguous (Table 3). The teachers who had read the leaflet or heard of it (49%) were largely persuaded by it. Physicians were split as to their opinion. However, the agreement figures among physicians are high (16/39) when one recalls that the leaflet dealt with health issues and spoke of E 330 (citric acid) as a major carcinogenic substance! Are opinion leaders themselves less knowledgeable than expected? To test this hypothesis, this particular additive was presented to the teachers and physicians under both its full name and, separately, its code name (E 330). Interviewees were asked if it was harmful or harmless (Table 4).

In general, when presented under its code name, citric acid is not recognized. Most teachers and physicians refuse to state an opinion:

Table 3. Impact of the Leaflet on Opinion Leaders

	Physicians (N = 101)	Primary School Teachers (N = 100)
Exposure[a]		
Have read the leaflet	18%	30%
Have heard of it	21%	19%
Have not read or heard	61%	51%
	Physicians Exposed to the Leaflet (N = 39)[b]	**Teachers Exposed to the Leaflet (N = 49)[b]**
Agreement with the leaflet[c]		
Agree strongly	4	17
Agree	12	22
Disagree	9	1
Disagree strongly	5	7
Check of the validity of the leaflet[d]		
No	31	38
Yes	8	11

[a] "Have you read or heard of the 'leaflet of Villejuif,' a list evaluating the toxicity of food additives and accusing some brands of marketing food products containing these dangerous additives?"

[b] Because of the small sample size, only raw figures are reported.

[c] (Asked of those exposed to the leaflet) "On the whole do you agree or disagree with the content of the 'leaflet of Villejuif'?"

[d] (Asked of those exposed to the leaflet) "Have you tried to check the validity of the content of the 'leaflet of Villejuif'?"

they do not know. Under its full name, only a minority of physicians and teachers say they do not know. Hence one of the roots of the leaflet persuasiveness resides in the fact that it presents only code names, thus preventing any recognition of the additives behind these codes.

However, it is striking that in spite of a clear identification, 3% of the physicians still declare "citric acid" as carcinogenic and 11% as bad for health! Among teachers, the leaflet of Villejuif is clearly the source of the carcinogenic image of E 330. Of the whole sample of teachers, 9% declared E 330 carcinogenic and 20% said it was bad for health. But among only those who had either read or heard of the leaflet (49 out of

Table 4. Percentage of Physicians and Teachers Declaring an Innocuous Additive to Be Carcinogenic or Bad for Health

	Physicians		Teachers	
	E 330	Citric Acid	E 330	Citric Acid
Carcinogenic	2%	3%	9%	1%
Bad for health	11	11	20	19
Innocuous	5	76	6	57
Do not know	82	10	65	23
N	101	101	100	100

QUESTION: I shall name a number of food chemical additives. Please tell me, for each of them, if it is carcinogenic, bad for health, or innocuous.

100), 16% declared E 330 carcinogenic and 18% said it was bad for health. The carcinogenic image of E 330 in the entire sample is exclusively due to those exposed to the leaflet. The similarity of "bad for health" figures suggests that teachers share on the whole the same suspicion of the food additives when they appear under their mysterious code name.

Finally, among both physicians and teachers exposed to the leaflet, only a minority has tried to check the validity of the leaflet (Table 3). These opinion leaders made up their opinion alone. They took the leaflet's word or just disbelieved it. Analysis of verbatims shows that no teacher called or wrote to the Hospital of Villejuif: the main modes of validity checking were speaking about it with colleagues or asking a local consumer organization. Among physicians, only two (out of eight who checked) directly asked the presumed source, the Hospital of Villejuif.

Despite their status and role as opinion leaders, physicians and teachers reacted in the same fashion as the layman; they rarely check the rumor's validity. When they do, instead of directly asking the source, they discuss the rumor within their reference group. This is clearly one of the motors of rumor diffusion (Kapferer, 1987).

In a postindustrial society, people have less and less understanding of their environment (Miller, Suchner, and Voelker, 1980). They rely on experts to make up their minds. Since many physicians themselves believed that citric acid was carcinogenic, taking a leaflet's word, one cannot expect the layman to be more resistant to persuasion.

THE LEAFLET'S CHANNELS OF DISTRIBUTION

How and where did people receive the leaflet? The national survey of October 1983 shows a large variety of modes of diffusion. Among all the respondents ($N = 939$), 3% found the leaflet in their mailbox, or someone handed it out at school (1%) or at the entry of banks, supermarkets, factories, and offices (5%). Not surprisingly it was also handed out in hospitals, or by nurses, while making visits to patients (2%). It was found posted in youth clubs, cultural associations, and sports clubs (1%). In fact, fooled by the leaflet's apparent credibility, ordinary people benevolently photocopied the leaflet and organized its diffusion in public places or among their friends, relatives, colleagues, and neighbors.

By far the largest quoted source of diffusion is the press (11%). Actually, the companies concerned by the leaflet hold files of the journals that reproduced the leaflet to warn their readers about food additives and some well-known brands. These are not the regional or national journals or magazines but the countless local, specialized, in-house association bulletins and magazines. The unqualified presence of the leaflet in this press not only diffuses the information at once to large audiences but confers additional credibility to the leaflet: it is akin to an endorsement effect.

When warned early enough, corporations attacked by the leaflet have prosecuted these journals on the basis of diffusion of false information as well as moral and financial prejudice. Arrangements were possible if the journal agreed to apologize in a subsequent issue and to distribute open denials to their readers.

After being simply handed out and later found in the press, the leaflet has recently appeared in books. Two cases were reported and led to lawsuits against the editors. The leaflet's contents were reproduced in a textbook on natural sciences for primary school pupils and in a vulgarization book on cancer written by a physician—a nonspecialist—promoting homeopathy. In both cases, the editors were required to issue denials and to make due corrections, but only in the next printings or editions of the books.

BEHAVIORAL EFFECTS OF THE RUMOR

A major concern of all companies having their brands quoted on the leaflet was the call to stop buying these brands. The author interviewed their brand managers; no figure could be obtained concerning the turnover losses due to the leaflet. Sales variations are caused by a number of variables; it is difficult to isolate the impact of the leaflet of Villejuif.

Table 5. Behavioral Effects of the Leaflet: Percentage of Respondents Who Reported Each Behavior

I posted it in my home	12%
I use it when going shopping	20
I have spoken about it to other persons	40
I have shown it to other persons	24
I xeroxed it for distribution	3
I gave it to somebody to be rediffused	2
None of these behaviors	39
N	121

QUESTION: Here are a number of reactions one can have to the leaflet. Could you tell me, for each of them, if it is what you did after you read the leaflet. (Asked of those recalling the leaflet; multiple answers were possible.)

In any case, no sharp drop in consumer sales which could be attributed to a boycott was ever noticed. The most anxious were two companies experiencing a slow but steady downward sales trend for products attacked by the leaflet: a processed cheese and an orange juice. However, the causal effect of the leaflet was clear-cut only on the shipments made directly to schools. Many schools stopped providing this brand at meals, explicitly out of fear of doing harm to children.

In order to assess the reality of the leaflet's persuasive impact on consumers' purchases, it was decided to conduct a field quasi-experiment. Five hundred copies of the leaflet (Figure 1) were distributed overnight in the mailboxes of houses in the middle-class suburbs of Rennes, a French provincial town. One week after the distribution, 150 housewives were contacted at their home by trained interviewers in order to measure the impact of the leaflet and also to give them all official denials of the leaflet, at the end of the interview. Among these, 81% (121 persons) spontaneously declared having recently read a leaflet warning consumers against additives in food products.

Table 5 summarizes the behavioral effects of the rumor. Interviewees were asked if they had followed one or more of six suggested actions. The main behavioral outcome is the communication of the rumor to other people, either by reproducing the leaflet or by speaking of its content. It was brought to stores to guide shopping by 20% of the leaflet readers.

Asked whether they had stopped buying specific brands or products after reading the leaflet, 19% (base = 121) admitted such an incidence. Purchase cessation intentions reach 69%; the majority of the remaining

31% declared that they did not usually buy the brands pointed out by the leaflet.

Spontaneous recall of the brands quoted in the leaflet proved high. For instance, among the people having read the leaflet (121) 45% recalled that Coca-Cola was said to be dangerous for health, 40% recalled the Banga orange juice, and 37% remembered the brand of cream cheese. These highest figures concern three brands which are very popular among children.

Finally, when asked to spontaneously recall which chemical food additives were carcinogenic, 29% of the leaflet readers said E 330, the only additive recalled among all those mentioned in the leaflet. Because it was singled out by the leaflet as the most dangerous, the readers polarized their attention on it.

Analysis and Discussion

THE LEAFLET PERSUASIVENESS

What makes a simple leaflet typed on an ordinary typewriter such a persuasive means of communication, not only among consumers but also teachers and physicians? Social psychologists working on persuasive communications have pointed out the importance of source, media, reception, and motivational factors in persuasion.

In fact, the leaflet of Villejuif optimizes each of these persuasion-enhancing variables. First of all, the leaflet is backed by a potent source effect. It presents itself as a direct release of research of this renowned center for cancer research, highly ranked on the two factors that make up source credibility: expertise and disinterest. Persuasion research has long demonstrated that high and low credibility sources lead to less remembering of the message content (McGuire, 1969:198). Acting as a "lazy organism," readers pay attention to the arguments only when source attribution is ambiguous. When confronted with a message backed up by an expert, they remember the source and the conclusion of the message. This explains why even physicians were persuaded by the leaflet and did not check its validity.

The source effect at work in the leaflet of Villejuif is not an exception. Most rumors rely on some kind of source effect. When they pass on a rumor, in order to enhance its credibility, people are prone to invent or to attribute hearsay to a close friend who knows some V.I.P. or reliable informant who confidentially told him some secret. Structurally this initial witness is granted the characteristics that make him an unquestionable source on the topic: health rumors mention a physician, crime rumors mention a policeman, a lawyer, or a judge. This is clearly an attribution process in order to ensure belief. In oral rumors, this process is difficult to catch. Based on a leaflet, the Villejuif rumor offered a unique opportunity to observe this process of creation of an unquestionable credibility. The leaflet started with an anonymous identification (a hospital of Paris) and ended as "Information communicated by the Hospital of Villejuif."

Unlike most rumors, this one circulates on paper because of the quantity of information it has to convey. The poor layout of the leaflet, typed on ordinary typewriters, may actually have contributed to the leaflet's persuasiveness: such leaflets are the typical media of underground resistance. There is a flavor of "David versus Goliath" in the leaflet of Villejuif (Fine, 1985). It presents itself as a defense of people's health obliged to use the underground to reach its public against large and powerful corporations accused of selling poisonous brands to make a profit. Spreading, diffusing, and copying the leaflet is motivated by the desire to alert one's friends, colleagues, and relatives about a danger. The whole process is akin to a crusade.

The fuel of this crusade is found in the latent anxieties stimulated by the permanent advances of science and technology in all areas of life (Sapolsky, 1986) and especially in food. Food has always been a very sensitive area, prone to stimulate rumors; most modern urban legends deal with contamination fears (Brunvand, 1981, 1984).

Actually, the leaflet fights against the steady rise of industrial food (unsurprisingly, the leaflet was often found at the cash registers of macrobiotic food stores). But it also has a political facet, alleging some kind of plot against the people's health, with the passive cooperation of the administration. The leaflet denounces the large corporations, guilty of advertising unhealthy brands to children, and the indifference (if not collusion) of the French Food and Drug Administration. This political facet of the rumor partly explains its persuasiveness among teachers, on the whole rather leftist in France. In the case of the "rumor of Orléans," E. Morin (1969) already found that teachers were influenced by the rumor because of its underlying political facet.

In addition to these motivational factors, the leaflet's persuasiveness was facilitated by the ambiguity of the code names of the chemical food additives. Pioneer rumor researchers (Allport and Postman, 1947; Shibutani, 1966) pointed out that rumors develop when a stimulus is both important and ambiguous. Rumor is viewed as a process of collective deliberation to provide meaning to the ambiguous stimulus. Since then, research has shown that this rumor formula clearly does not explain all types of rumors (Rosnow, 1980; Kapferer, 1987). For instance, some rumors seem to emerge from scratch: one cannot find any event or fact which could have provoked word of mouth. The "vanishing hitchhiker" (Brunvand, 1981) or the "rumor of Orléans (Morin, 1969) are typically such rumors. They are mostly phantasmatic and do not need any reality to burst out locally.

If the basic rumor formula certainly does not account for all types of rumors, it provides a model for understanding the leaflet of Villejuif. There is no doubt that cancer, poison, and food are extremely important and involving topics, especially if they concern one's children. It is significant that chemical food additives have been designated by codes whose meanings are ambiguous. As all codes, they suggest some need for secrecy, of maintaining their object far from the public's understanding, as if some danger were to be hidden.

THE ANTIRUMOR DIFFICULTIES

The success of the leaflet of Villejuif is also a result of the weakness of antirumor attempts. First, corporate policies made collective action impossible. Although the leaflet attacked the brands of the largest food companies operating in France, they did not unite to react. For instance, the Coca-Cola Corporation decided to stay mute for fear of seeing a more important question being raised: the case of sugar. What would have been easy to do, if done rapidly and forcefully by the pooling of financial resources, became too costly for the few corporations willing to attack the rumor.

Due to the absence of a common strategy, a low-profile approach was followed. Official sources were urged to release denials and reassuring communiqués. Indeed, the French Ministry of Health, the Villejuif Hospital, cancer specialists, and even consumer organizations all made unambiguous denials. But these denials rarely received significant consumer exposure. Typically the press would publish them once and consider any repetition unnecessary.

Formal police investigations were undertaken in an attempt to identify the source of the leaflet, with no result. After a few months, the process of leaflet diffusion had become autonomous.

When it was known that leaflets were being distributed at a specific place, lawsuits were undertaken, but the judiciary process was slow and lenient. Leaflet distributors claimed to have been fooled by the apparent credibility of the leaflet and the reference to the Villejuif Hospital. Judges were more severe about media publications of the leaflet: denials had to be clearly and openly published at their expense in the daily or magazine. Coming too late, in a dispersed and scattered way, these actions hardly deserve to be called antirumor actions. In practical terms this is equivalent to just waiting for the natural end of the rumor. Through time, consumer fears should diminish and the leaflet's fear appeal lose its effectiveness. Certainly, the leaflet's diffusion pace is presently slower in France. But it is rapid in Italy, whose consumers have just started being poisoned—by the rumor.

In retrospect, the involved companies first followed a "do nothing" strategy and later on tried the "do something locally" and "do something discreetly" strategies. They refused to frontally attack the rumor by using all the media resources available. This attitude is widespread. F. Koenig (1985:167) also pointed out that "there appears to be a general, almost mystical reluctance on the part of some public relations people to confront a rumor problem directly. There seems to be a lurking fear that an open campaign will add fuel to the fire." If the rumor is a real problem, one cannot afford to wait until all else fails. Meanwhile, one leaves the monopoly of speech to the rumor. The silence of the involved companies made the leaflet of Villejuif all the more persuasive.

THE LEAFLET AS RUMOR

In this article, the leaflet of Villejuif was called a rumor. Since some may question whether this leaflet is a rumor, theoretical arguments are presented now, after the reader is acquainted with the account of the dissemination and acceptance of a proposition rejected by specialists and officials.

Classifying the leaflet of Villejuif as a rumor supposes an empirical definition of rumor. The French sociologist Edgar Morin (1969) defined "pure rumor" as topical information circulating in a social group, off media and largely unfounded. The leaflet meets these criteria: it is topical, concerning existing brands actually purchased by consumers; it circulates in specific segments of the population; its circulation was initially restricted to being passed from hand to hand—only later did it appear in the "small press," but still off media; it is disavowed by all cancer researchers and toxicologists.

The behavior and characteristics of the leaflet are those expected of any rumor (Rosnow and Fine, 1976). It has unknown origin. It makes use of an alleged unquestionable source. It polarizes the readers' attention on a few specific points (one additive, a few well-known brands). Finally, there is distortion since the real and official facts concerning the food additives have been either exaggerated or changed.

On the other hand, one classical evolution of oral rumor—sharpening—did not take place. Being couched on a leaflet, the full original message was always available and did not have to undergo the sharpening effects of selective forgetting and remembering.

However, the leaflet of Villejuif is not just any kind of rumor. Since the very first leaflet had to be typed by someone, it is a planted rumor, which resulted in public disinformation. In modern terms, some would call it a planned disinformation by means of rumor (Lienhardt, 1975). Once the first leaflet was distributed, the spontaneous social dynamics of rumor have ensured its amazing nationwide diffusion. It is somewhat perplexing to see how, once planted, despite official denials, false information circulates and gains such a high degree of popular acceptance.

References

Allport, G., and L. Postman (1947)
The Psychology of Rumor. New York: Holt.
Brunvand, J. H. (1981)
The Vanishing Hitchhiker. London: Pan Books.
——— (1984)
The Choking Doberman. New York: Norton.
Cantril, H. (1966)
The Invasion from Mars. New York: Harper Torchbooks.
Degh, L., and A. Vazsonyi (1975)
"The hypothesis of multi-conduit transmission." Pp. 207–252 in D. Ben-Amos and K. Goldstein (eds.), Folklore, Performance, and Communication. The Hague: Mouton.
Fine, G. A. (1985)
"The Goliath effect." Journal of American Folklore 98:63–84.
Kapferer, J. N. (1987)
Rumeurs: Le plus vieux média du monde. Paris: Le Seuil Editions. English translation as Rumors: World's Oldest Media. New Brunswick, NJ: Transaction Books, forthcoming.
Koenig, F. (1985)
Rumor in the Marketplace. Dover, MA: Auburn House.
Lienhardt, P. G. (1975)
"Interpretation of rumor." Pp. 105–131 in Social Anthropology: Essays in Memory of Sir Evans Pritchard. Oxford: Oxford University Press.
McGuire, W. J. (1969)
'The nature of attitudes and attitude change." Pp. 136–314 in G. Lindzey and E. Aronson (eds.), Handbook of Social Psychology, vol. 3, 2d ed. Reading, MA: Addison-Wesley.
Miller, J. D., R. W. Suchner, and A. M. Voelker (1980)
Citizenship in an Age of Science. New York: Pergamon.
Morin, E. (1969)
La rumeur d'Orléans. Paris: Le Seuil Editions.
Rosnow, R. (1980)
"Psychology of rumor reconsidered." Psychological Bulletin 87:378–391.
Rosnow, R., and G. Fine (1976)
Rumor and Gossip: The Social Psychology of Hearsay. New York: Elsevier.
Sapolsky, H. M. (1986)
Consuming Fears. New York: Basic Books.
Shibutani, T. (1966)
Improvised News. New York: Bobbs Merrill.

FROM RACISM TO TOKENISM
THE CHANGING FACE OF BLACKS IN *NEW YORKER* CARTOONS

RUTH THIBODEAU

Abstract Cartoons published in the weekly magazine *New Yorker* during the years 1946–87 were analyzed to determine how often blacks appeared as characters and whether the proportion of blacks portrayed had changed over time. Cartoons were further examined for possible changes in the styles of characterization of U.S. blacks since the post–Word War II period. Although black characters were extremely rare throughout the 42 years, the relative numbers of blacks depicted decreased with the passage of each time period. Styles of portraying black characters also changed markedly over the years, in keeping with the changing status of blacks in the United States. All cartoons from the earliest period presented U.S. blacks in stereotypic occupational roles, cartoons in the late 1960s and early 1970s were dominated by racial themes, and blacks appeared in "token" roles in the majority of cartoons from the most recent period investigated. In the entire 42 years, only a single U.S. black appeared as a main character in a cartoon in which race was completely irrelevant.

"Laughter is a weapon against bigotry," Gordon Allport wrote in his classic 1954 treatment of prejudice. "It too often lies rusty while reformers grow unnecessarily solemn and heavy-handed" (Allport, 1954:509). Yet humor can cut both ways: used insensitively or in the form of racial or ethnic put-downs of one group by another, it often serves to solidify in-group ties and harden hostile attitudes toward out-group members (Burma, 1946; Martineau, 1972). Given its double-edged nature, humor must be handled judiciously by those wishing to heal rather than deepen existing racial wounds.

RUTH THIBODEAU is a Ph.D. candidate in Social Psychology at the University of California, Santa Cruz. The author is indebted to Elliot Aronson and Thomas F. Pettigrew for their valuable suggestions on earlier drafts of this article. Special thanks are also due to Richard Gendron, who helped with data collection, and Alysia Parker and Tim Tumlin, who served as coders. Correspondence concerning this article should be addressed to Ruth Thibodeau, Psychology Board of Studies, Adlai E. Stevenson College, University of California, Santa Cruz, CA 95064.

Public Opinion Quarterly Volume 53:482–494 © 1989 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/89/0053-04/$2.50

Apparently mindful of such problems, mainstream cartoonists who wished to bring blacks into their work in the late 1960s voiced concern about how their attempt to make a positive social statement would be greeted. For example, Charles Schulz, creator of the "Peanuts" comic strip, introduced the black character "Franklin" in 1968, only after considerable soul-searching and self-doubt: "Up until now I had always come to the conclusion that I wasn't capable of doing this thing properly. I had always thought that Negro people might feel I was doing it in a patronizing or condescending manner. Then, I received two letters from Negro fathers asking that Peanuts be integrated. That was the turning point" (Newsweek, 1968:67).

Given the turbulent social climate of the late 1960s and early 1970s, signs of concern by mainstream cartoonists over how best to treat racial themes, or merely black faces, in their work seems hardly surprising. Now that much of the furor of those years has quieted down, how have blacks been handled in cartoon humor in the wake of the civil rights movement? Or, for that matter, how were blacks depicted in cartoons prior to this eventful period in U.S. black–white relations? Finally, have black characters appeared with increasing frequency over the years, particularly in publications known for sophisticated, liberal editorial policies?

Given the increased representation of blacks on television since the 1960s (Baptista-Fernandez and Greenberg, 1980), one might expect to find few black characters in cartoons published during the pre–civil rights years, followed by a sharp increase in the late 1960s and early 1970s, with this increase perhaps leveling off in subsequent years up to the present. On closer scrutiny, however, these predictions seem less plausible. In the first place, prior to the civil rights movement, a cartoonist could probably portray blacks in an openly stereotypic or derogatory fashion with relative impunity. In particular, the depiction of blacks in subservient roles, under conditions of segregation, or as "savages" in jungle settings would have raised few eyebrows among white viewers in the 1940s and early 1950s. On the other hand, antiracist norms (particularly, prohibitions against stereotyping blacks) that grew out of the civil rights and black militancy movements clearly made some cartoonists (e.g., Charles Schulz) more self-conscious about how blacks and racial issues should be treated in their work. If racial issues were tough to tackle in the volatile 1960s and early 1970s—when the times virtually cried out for their inclusion in all forms of media—there is little reason to expect more blacks to appear in cartoons in the years following the civil rights movement, now that the issue of the condition of blacks in America has, by all outward signs, taken a back seat to other concerns.

Finally, without pressing racial issues serving as a strong motivation to portray blacks, cartoonists in recent years *could* exercise the option of depicting a black character prominently in a cartoon for nonracial reasons, as an "Everyman" representing all humanity rather than blacks as a special subgroup. To do so, however, might cause a different, but equally problematic, form of misinterpretation to rear its head: viewers of such cartoons might seach for racial meanings where none were intended. Thus, for all these reasons, an investigation of cartoons published before and after the civil rights years should reveal a decrease in the appearance of black characters over time, as well as a marked transformation in the forms taken by such portrayals.

Citing humor as an especially sensitive indicator of social attitudes, researchers have investigated cartoon portrayals of women (Meyer et al., 1980; Kramer, 1974), the elderly (Smith, 1979), businessmen (Cullather, 1983), sex-role stereotypes (Anderson and Jolly, 1977), and sexual violence (Malamuth and Spinner, 1980). However, the investigation of the portrayal of blacks in cartoons is a topic that has been largely ignored, save for a single study by Greenberg and Kahn (1970), which looked at the treatment of blacks in *Playboy* cartoons in all issues of 1956, 1958, 1960, 1964 to 1968, and the first seven months of 1969. These researchers found that, until 1967, contemporary blacks were seen in less than 1% of *Playboy* cartoons. Analysis of succeeding years revealed an abrupt increase in the presence of blacks to nearly 4% of cartoons in 1969. These results are consistent with *Playboy*'s editorial policy during the late 1960s, which called for the active solicitation of cartoons involving "integrated themes" (Newsweek, 1968). Among their findings, Greenberg and Kahn observed that in 1968–69, some 75% of the cartoons containing black characters involved racially motivated humor. They also found that blacks were rarely found as main characters in cartoons involving sex-related humor.

In an effort to capture possible changes in the depiction of blacks over a wider time span, the present study looks at cartoons published from the post–World War II years predating the civil rights movement up through the mid 1980s. Cartoons appearing in the weekly magazine *New Yorker*, in continuous circulation since 1925, were selected for this investigation for several reasons. In quality as well as sheer quantity, *New Yorker* cartoons are considered unsurpassed. Celebrated for decades as the "font of American humor," (Kanfer, 1975:56), the *New Yorker* every week considers some 2,500 cartoon ideas and in each issue publishes roughly 15 to 20 cartoons drawn by the nation's best cartoonists (Feinberg, 1981). Widely circulated beyond the magazine's pages, these cartoons are considered a "means of gauging the changing winds of modern society. These cartoons adorn countless refrigerators, offices, and bulletin boards. For millions of people they are a source of dependable weekly pleasure" (Feinberg, 1981:54). In addition, the *New Yorker*'s impressive standards of journalism, and its reputation as "one of the most powerful cultural institutions in the country" (Unger, 1983:29) have attracted over a half million sophisticated, affluent subscribers. Finally, the magazine's left-leaning sympathies on important social issues and its insistence on "good taste" in the work of its cartoonists (Newsweek, 1968; Feinberg, 1981) suggest that portrayals of blacks in *New Yorker* cartoons could be expected to reflect the sensibilities of well-educated, liberal whites in the United States.

The present investigation examined cartoons published in the *New Yorker* during the years 1946–87 to determine how often blacks appeared as characters and whether the proportion of blacks portrayed had changed with the passage of time. Cartoons were further analyzed for possible changes in the styles of characterization of U.S. blacks since the post–World War II period. It was hypothesized that the percentage of all black characters would decline over time, with the depiction of non-U.S. blacks in aboriginal settings accounting for the relatively higher representation of blacks during the first two decades studied (1946–65). It was also predicted that portrayals of U.S. blacks during the earlier years of the 42-year period would be primarily of a stereotypic nature.

Method

Every issue of the *New Yorker* published during the years 1946–87 (inclusive) was examined, yielding a total 2,193 issues. The frequency of all cartoons in each issue was tallied, with a separate count also maintained for cartoons depicting black characters. Cartoons with blacks were than categorized as containing either U.S. blacks or non-U.S. blacks. Those cartoons designated as containing U.S. blacks were subdivided into active vs. passive portrayals. To be categorized as active, the character was required to have a speaking role or, in an uncaptioned cartoon, the character's action or perspective had to be the central focus of the cartoon. The subsample of cartoons with U.S. blacks was further analyzed using the following four "portrayal" categories, which were mutually exclusive and hierarchically coded:

1. Race-related humor. The issue of race is central and the character is black for a specifically racial reason, most often representing blacks as a group.
2. Stereotyped role. The character is engaged in an occupation or behavior stereotypically associated with blacks. Typically, the black character is shown in an inferior-status occupational role

Table 1. Frequency of Blacks Portrayed in *New Yorker* Cartoons

	Frequency		
Time Period	Total Cartoons	All Blacks	U.S. Blacks
1946–55	7,364	101 (1.37%)	14 (0.19%)
1956–65	7,659	87 (1.14%)	9 (0.12%)
1966–75	8,611	77 (0.89%)	62 (0.72%)
1976–87	12,240	58 (0.47%)	51 (0.42%)
All years	35,874	323 (0.90%)	136 (0.38%)

relative to whites or—in recent years especially—is shown in such stereotypic roles as professional athlete or entertainer.

3. Token role. The character appears among a crowd, group, or pair of role-similar others in a racially integrated setting. The main character of the cartoon is not black.
4. Non-race-related theme. The main character is black, and there is no obvious racial theme in the cartoon. The character could be replaced by a white character without altering the cartoon's meaning.

Coding was done by two advanced undergraduate psychology majors, one a black woman and the other a white man. As calculated using Scott's pi (Scott, 1955), which corrects for agreement by chance, intercoder reliability was 93.5% for U.S. versus non-U.S. blacks, 91.6% for active versus passive, and 81.7% for the "portrayal" categories.

Results

The data were grouped into three 10-year periods (1946–55, 1956–65, 1966–75) and one 12-year period (1976–87) to facilitate comparisons of the proportions of blacks portrayed and the differences in styles of characterization over the 42 years.

As is evident in Table 1, blacks appeared extremely infrequently in cartoons published in the *New Yorker* throughout the 42 years, with a meager high of 1.37% during 1946–55, to an almost negligible 0.47% in 1976–87. Despite such low frequencies, the proporton of blacks depicted differed significantly over the years, as revealed by chi-square analysis comparing all four periods ($\chi^2 = 48.04$, $df = 3$, $p < .0001$). In line with this study's predictions, the proportion of black characters

Table 2. Portrayal of U.S. Blacks in *New Yorker* Cartoons

	Time Period			
Portrayal	1946–55 ($N = 14$)	1956–65 ($N = 9$)	1966–75 ($N = 62$)	1976–87 ($N = 51$)
Active	4 (28.6%)	3 (33.3%)	14 (22.6%)	3 (5.9%)
Passive	10 (71.4%)	6 (66.6%)	48 (77.4%)	48 (94.1%)
Race-related	0	4 (44.4%)	26 (42%)	1 (1.5%)
Stereotyped	14 (100%)	4 (44.4%)	11 (17.7%)	11 (21.6%)
Token	0	1 (11.2%)	25 (40.3%)	38 (74.5%)
Non-race-related	0	0	0	1 (1.5%)

appearing in *New Yorker* cartoons decreased over time, with more blacks appearing in the years 1946–65 than in 1966–87 ($\chi^2 = 35.697$, $df = 1$, $p < .0001$).

As was also predicted, the portrayal of non-U.S. blacks (84% of which were blacks depicted in aboriginal settings) accounted for the higher proportions of all blacks appearing in cartoons during the earliest two decades studied (1946–65). In fact, when only U.S. blacks are considered, significantly more black characters appeared in cartoons during 1966–87 than in the two earlier decades, 1946–65 ($\chi^2 = 34.34$, $df = 1$, $p < .0001$). It is worth noting, however, that the proportion of U.S. black characters appearing in the most recent period, 1976–87, was significantly lower than in 1966–75 ($\chi^2 = 8.63$, $df = 1$, $p < .005$).

When cartoons were analyzed for active versus passive characterization, it was found that U.S. blacks have most often been portrayed in passive roles during the entire 42-year period (Table 2). Over time, however, this form of depiction has become more prevalent, rising from 71.4% in 1946–55 to 94.1% in 1976–87 ($\chi^2 = 3.76$, $df = 1$, $p < .05$).

The prediction that U.S. blacks would appear primarily in stereotyped roles during the earlier years of the 42-yer period was also confirmed. Indeed, all 14 cartoons seen in 1946–55 presented U.S. blacks in a stereotyped manner, most often in subservient occupational roles (e.g., domestic servant). Comparing the first two and last two periods, stereotypic portrayals of U.S. blacks fell from 78.3% in 1946–65 to 19.5% in 1966–87 ($\chi^2 = 29.05$, $df = 1$, $p < .0001$).

Characterization of U.S. blacks in token roles—usually portrayed amidst a group or crowd of whites, often in the distant background—was not in evidence in the 1946–55 decade and was seen only once during 1956–65. In later years, however, the percentage of U.S. blacks depicted as "tokens" rose from 40.3% in 1966–75 to become at 74.5% the predominant form of portrayal in the 1976–87 period ($\chi^2 = 13.26$, $df = 1$, $p < .0001$).

While no cartoons with race-related themes were seen in the earliest decade, four of the nine cartoons with U.S. blacks in the 1956–65 period involved a racial message. Not surprisingly, cartoons depicting race-related humor were most prevalent in 1966–75, accounting for 26 of the 62 cartoons (42%) with U.S. blacks published during these years. These topical cartoons did not appear with equal frequency throughout the decade. Rather, 85% of the race-related cartoons seen in 1966–75 were concentrated in the years 1967–70 during the height of public attention to the issue of race relations. In contrast with the 1966–75 period, in 1976–87 only one cartoon of the 51 depicting U.S. blacks involved a racial theme ($\chi^2 = 22.44$, $df = 1$, $p < .0001$).

To qualify for the category "non-race-related theme," a cartoon's thematic content had to be unrelated to race, and the black character had to be depicted neither stereotypically nor as a "token." Only one cartoon analyzed in this study was able to meet these requirements. In other words, only a single U.S. black appeared as a main character in a cartoon in which race was irrelevant during the entire 42 years investigated.

Discussion

Readers of the *New Yorker* over the last 42 years have had little occasion to encounter cartoons with black characters—particularly U.S. blacks, which have appeared in less than 1% of cartoons published since World War II. While the number of U.S. blacks depicted has remained negligible, the particular styles of characterization have changed considerably during the 42-year period, often reflecting historical alterations in the status of blacks, as well as white attitudes toward blacks in the United States. In the following discussion, transformations in the portrayal of blacks in *New Yorker* cartoons will be reviewed.

1946–1955

Of the 101 cartoons containing black characters during the immediate postwar decade, only 14 portrayed U.S. blacks rather than African slaves or tribesmen. Of this handful, every cartoon presented U.S. blacks in stereotypic occupational roles such as domestic servants, bellhops, stable hands, etc. However, blacks occasionally emerged as "full-size" (as opposed to background) characters in these cartoons and at times played central, active roles in the humor, which was in some cases at the expense of a white character. For example, a cartoon published 17 September 1949 shows a white woman in formal riding attire sheepishly trudging back to the stables on foot, minus her horse. In the foreground, a black stable hand calls out to her: "Ah, there you are at last, Mrs. Jennings. Your horse got back at two-thirty, so we'll only charge you till then." This prominent, active style of portrayal may reflect the relative security many white liberals—such as the sophisticated, affluent readers of the *New Yorker*—presumably felt regarding their relationship to blacks during this period. With blacks safely relegated to inferior-status positions relative to whites, a black character could be allowed to poke fun at a white character occasionally, while posing little threat (psychological or otherwise) to the sanctity of white supremacy.

1956–1965

Given the rising importance of racial issues in the United States during the later years of the 1950s and the early 1960s, it is striking that this decade yielded the fewest cartoons depicting U.S. blacks. While 87 cartoons with aborginal and other non-U.S. blacks appeared during these years, only nine cartoons featured U.S. blacks. Still, these cartoons offer an intriguing look at the changing status of blacks during this period. There were, for example, just as many race-related cartoons as stereotypic ones, in contrast to the previous decade in which every cartoon presented blacks stereotypically. These stereotyped depictions differ qualitatively from those of earlier years, with only one out of the four cartoons showing a black in a lower-status occupational role and the remainder depicting blacks as professional athletes.

The style of cartoon most representative of this decade, however, typically shows a group of tribal Africans getting the best of white missionaries or explorers (often by cooking them up for dinner) or following a gun-toting white man through the jungle, carrying his belongings on their heads, native-style. A possible clue to why so few U.S. blacks appeared during this period may be seen in a cartoon published 18 April 1959 which depicts the typical jungle scene with the black-white roles reversed. Several white men trail behind a black man sporting a safari hat and rifle, with the whites now carrying the black man's possessions on *their* heads. Hidden from view and appraising the scene from afar, two white hunters share the comment, "My God, I didn't think that the pendulum had swung this far the other way."

Thus, one plausible explanation for the dearth of U.S. blacks portrayed in cartoons in the years 1956–65 may lie in the inability or reluctance of the white majority to accommodate the rapidly changing status of blacks during this period. In other words, it may have become increasingly difficult to depict blacks in subservient roles in the unselfconscious manner typical of the 1946–55 period, while the presentation of blacks as equals to whites was, as yet, largely unthinkable.

1966–1975

Cartoons involving race-related humor—with themes typically focusing on black–white relations, black militancy, and poverty—prevailed during this decade, accounting for 26 out of a total 62 cartoons depicting U.S. blacks. Only a few cartoons portrayed blacks in aboriginal settings and, in nearly every instance, the humor turned on the ludicrous ethnocentrism of whites, with the aboriginal characters depicted as active, intelligent, and sophisticated individuals. Gone are the frenzied hordes of spear-wielding natives. Blacks also appear in token roles in cartoons from this period, with black faces showing up, for example, in a crowd of student protesters, subway riders, or summer campers.

The years 1966–75 also mark the strong emergence of race-related cartoons in which U.S. blacks appear on an equal-status basis with whites in social settings. Although blacks are now seen in middle- and upper-middle-class settings as the social peers of whites, the absolute passivity of these otherwise dignified portrayals is striking. In these cartoons, not a single black spoke in a cartoon in which whites also appeared. When blacks were given speaking parts—and this was quite rare—they invariably spoke only to other blacks.

The passivity of blacks portrayed in the company of whites during this period should not be mistaken for their lack of importance in these cartoons. To the contrary, the presence of a black character is absolutely crucial to the humor which, in most cases, serves to point out the clumsy, patronizing efforts of whites to come to grips with equal-status relationships with blacks. For example, a cartoon published 8 June 1968 (a variant to several similar cartoons seen in the late 1960s) shows a solitary, stately black man conversing with a white man at an otherwise all-white cocktail party. The white man says to the black: 'I understand you're a very militant fellow. Sock it to me!" Sophisticated white viewers of such cartoons would probably have laughed at themselves—or more likely, other whites—while suffering minimal uneasiness regarding their own racial attitudes and behavior. On the other hand, a cartoon dealing more directly with the thorny issue of black–white relations would have perhaps been too unsettling to allow for a humorous response by most whites. It is far less threatening to see a black depicted as passively listening to a white make an ass out of himself than it would be to behold an angry black actively "socking it" to another white.

1976–1987

Cartoons featuring U.S. blacks in 1976–87 depart abruptly from the race-related humor that characterized the previous decade. Concern with black–white relations apparently no longer occupied a central role in the consciousness of *New Yorker* cartoonists and, presumably, their audience. With racial issues now on the back burner, how have blacks been characterized in cartoons from the most recent period investigated? Minimally, if at all. In this period, 74.5% of black characters appeared as "tokens" and, as in the previous decade, blacks in such cartoons are typically shown submerged in a group or crowd of whites and are usually found off in the background. As tokens, these characters *do* appear on an equal-status basis with role-similar whites, but here the parity consists solely in being equally insignificant.

Apart from "tokens," stereotypic portrayals of U.S. blacks constituted the next most common form of depiction in 1976–87. The stereotypes have changed dramatically, however, from the subservient occupational roles witnessed in the 1946–55 period. Occupational roles are still seen, but now they are of a far more glamorous nature: professional athletes, musicians, and entertainers. As we saw earlier, blacks portrayed in stereotypic, low-status positions were, at times, afforded an active, central role in the cartoon's humor. Not so now: blacks appearing in present-day, higher-status roles are, almost without exception, portrayed passively. In fact, these characters tend to be depicted as insignificantly as the "tokens" discussed above. At best, they have a tangential role in the humor and often appear in the distance, somewhat out of focus.

Given the minimal portrayal of blacks as "tokens" and in stereotypic roles, it is scarcely surprising that nearly all cartoons analyzed for this period fell into the passive category. This passivity is of a different, more thorough variety than that seen in the previous decade, when blacks had a vital, if strangely silent, presence in the service of race-related humor. With race no longer a central issue, blacks have simply faded into the background.

Blacks as "Everyman"?

As we have seen, the depiction of blacks in peripheral "token" roles has characterized *New Yorker* cartoons from the most recent period studied. What alternatives to this kind of minimal portrayal are open to the cartoonist? In what kind of cartoon could blacks be presented as central characters without stereotypes or racial issues to "justify" their presence? One form of pictorial humor—less dependent on particular stereotypes or highly defined roles—is the "human condition" type of cartoon which typically portrays an ordinary person apprehending something humorous in the environment or confronting the humbling limitations of being human. The sole purpose of the character in such a cartoon is to stand as an easily recognizable symbol for "Everyman": not a specific kind of human being, but a prototypic representation of generic humanity. The type of character usually seen in this "Everyman" role, at least in mainstream publications, is the white middle-class male.

The coding scheme used in the present study allowed for the possibility of detecting blacks in this "Everyman" role, yet only a single cartoon in 42 years met the requirements of the "non-race-related humor" category. Published 25 May 1987, the cartoon shows a car parked in a "no parking" zone, and standing next to it is a uniformed parking-control officer—a black woman. Hand poised to issue a parking ticket, she stares at the car in amazement, baffled by the odd sight of a sign on the dash that reads "MBA on call."

What makes this cartoon unique is the fact that the central character is black *for no particular reason*. The officer is the sole character in the cartoon, which has no obvious or subtle racial meaning. And, although the character doesn't speak, her perspective is clearly crucial to the cartoon's impact: it is through her eyes that the viewer perceives the potential humor in the situation. Why are there no other examples of such cartoons? Why aren't blacks being portrayed prominently for nonracial reasons?

Apart from outright racism, one possible explanation involves the desire of cartoonists or their editors to *avoid* appearing racist. Yet, it might be countered, if stereotypes and crudely racist portrayals are consciously avoided, this should not be a problem. However, this is exactly the dilemma: because norms regarding what in fact constitutes racism (particularly on a subtle level) may not be altogether clear in the post–civil rights era, cartoonists face the risk of inadvertently invoking a stereotype associated with blacks and, thus, may find themselves open to the charge of racism.[1]

Stereotypes, after all, are composed of traits and behaviors that, in and of themselves, are common to all human beings. As stereotypes, however, they are overgeneralized to a particular group and become destructive when they provoke prejudiced, undifferentiated attribu-

---

1. Indeed, the non-race-related cartoon described above may escape this potential pitfall largely through the depiction of the black character in a highly circumscribed fashion. In uniform and acting in an official capacity, the traffic officer appears in a role that is easily and instantly recognized. The fact of her race is subordinate to and overwhelmed by her official function as a civil servant, and the clearly defined nature of this role invites no misunderstanding as to the cartoon's meaning. More importantly, it invokes no unintended stereotype associated with blacks.

tions about the personalities of individuals belonging to the stereotyped group. Thus, the very human characteristics drawn upon by the cartoonist to portray *any* individual, regardless of race, are also likely at times to have strong associations with a particular group, such as blacks. For the white cartoonist wishing to portray blacks in nonracist ways, these tenacious associations are bound to be problematic. Given the salience of racial information, it is likely that a specifically racial interpretation will result when a black character is shown engaging in a *human* behavior—if that behavior is simultaneously a potentially stereotypic one—even when a racial meaning was not intended to be part of the humor. If this behavior is also unflattering (and cartoons often expose the less noble aspects of the human condition) then the possibility of the cartoon assuming a racist tone, at least in the eyes of some viewers, becomes even more probable.[2]

As an illustration, imagine a cartoonist depicting a black as the main character in a cartoon centering on the theme of everyday *human* laziness. If this cartoon appeared in the pages of a mainstream publication, rather than one published largely by and for black people, viewers might perceive it as intentionally delivering a stereotypic—and very racist—message about blacks specifically, rather than people in general.

One avenue for future research would be to explore more deeply some of the questions raised by this study. For example, if white viewers were exposed to cartoons with blacks depicted as main characters for no racial reason, would they find such cartoons equally funny as the same cartoons in which whites appear? Simply because the character is black, would most viewers look for racial interpretations of the humor? When none is forthcoming, would the viewer's puzzlement undermine the cartoon's impact? These issues await empirical investigation.

Finally, the present study examined cartoons from only one magazine, the *New Yorker*, which appeals to an unusually well educated and affluent readership. It is conceivable that magazines directed at a broader audience may publish more (or perhaps even fewer) cartoons with black characters, or they may portray black characters in roles different from those found in the *New Yorker*. Thus, additional research analyzing cartoons from a variety of magazines would be useful in determining the possible limitations of the present findings.

---

[2] While the fear of being suspected of racism might possibly inhibit the use of blacks in cartoons drawn by white artists, this presumably would be less of a problem for black cartoonists. As blacks, they would most likely be given the benefit of the doubt in the event that their depictions of fellow blacks seemed to skirt the edge of a possibly racist interpretation. Rather, the dilemma faced by black cartoonists would seem to lie more fundamentally in getting their work published in mainstream magazines and newspapers (Brandon, 1981). Also, it should be noted that comic strips—unlike the one-panel cartoons investigated here—allow for a less ambiguous interpretation of a cartoonist's efforts, since strips provide more time and space for character and plot development. Characters in comic strips usually have a continuous identity, a "history" developed from day to day or week to week, thus rendering unintended interpretations of their words or action less likely. Further exploration of these issues is, however, beyond the scope of this study.

References

Allport, G. W. (1954)
The Nature of Prejudice. 25th anniversary ed. Reading, MA: Addison Wesley.
Anderson, R. E., and E. Jolly (1977)
"Stereotyped traits and sex roles in humorous drawings." Communication Research 4:453–484.
Baptista-Fernandez, P., and B. S. Greenberg (1980)
"The context, characteristics, and behaviors of blacks on television." Pp. 13–21 in B. S. Greenberg (ed.), Life on Television: Content Analyses of U.S. TV Drama. Norwood, NJ: Ablex
Brandon, B., Jr. (1981)
"The serious side of cartooning." Crisis 88:476–478.
Burma, J. H. (1946)
"Humor as technique in race conflict." American Sociological Review 11:710–715.
Cullather, J. L. (1983)
"Has the laughter died? Musings on *The New Yorker*'s business ethics cartoons." Business Horizons 26(2):30–33.
Feinberg, A. (1981)
"Still at the old drawing board." Horizon, January, 54–61.
Greenberg, B., and S. Kahn (1970)
"Blacks in *Playboy* cartoons." Journalism Quarterly 47:557–560.
Kanfer, S. (1975)
"*The New Yorker* turns fifty." Time, 3 March, 56–57.
Kramer, C. (1974)
"Wishy washy mommy talk." Psychology Today, June, 82–85.
Malamuth, N. M., and B. Spinner (1980)
"A longitudinal content analysis of sexual violence in best-selling erotic magazines." Journal of Sex Research 16:226–237.
Martineau, W. H. (1972)
"A model of the social functions of humor." Pp. 101–125 in J. H. Goldstein and P. E. McGhee (eds.), The Psychology of Humor: Theoretical Perspectives and Empirical Issues. New York: Academic Press.
Meyer, K., J. Seidler, T. Curry, and A. Aveni (1980)
"Women in July fourth cartoons: A 100-year look." Journal of Communication 30(1):21–30.
Newsweek (1968)
"Charlie Brown's new pal." Newsweek, 29 July, 66–67.
Scott, W. A. (1955)
"Reliability of content analysis: The case of nominal scale coding." Public Opinion Quarterly 19:321–325.
Smith, M. D. (1979)
"The portrayal of elders in magazine cartoons." Gerontologist 19:408–412.
Unger, C. (1983)
"Murmurings at *The New Yorker*." New York, 28 November, 28–40.

563

THE POLLS—A REVIEW
THE 1988 PRESIDENTIAL PRIMARIES

LARRY M. BARTELS AND
C. ANTHONY BROH

Pollsters have been involved in presidential elections for more than half a century, and by now the best are able to gauge public reactions to the competing candidates with impressive accuracy. But polling in state presidential primaries is, by comparison with that in general election campaigns, both a newer and a more difficult art: newer because the contemporary system of presidential nomination, with its intensive public focus on the candidates' performances in a long and decisive series of weekly primary elections, is less than twenty years old; more difficult because the primary season involves a more complex set of contests than the subsequent general election campaign. Primary season involves more candidates, fewer voters, a less-knowledgeable and less-committed public, and less-understood campaign and election rules than the general election. In 1988 Jesse Jackson, the first black candidate to have a chance to win the nomination of a major political party, introduced additional complexity to the primary season. By generating an essentially new political situation, Jackson helped to illuminate both the pollsters' standard operating procedures and their occasional innovations.

How—and how well—do pollsters deal with these primary season difficulties? Here we describe and assess the efforts of three prominent polling organizations in the 1988 presidential primary campaign: the Gallup Organization, the Harris Poll, and the CBS News/*New York*

LARRY M. BARTELS is Associate Professor of Political Science at the University of Rochester. C. ANTHONY BROH is the Registrar of Princeton University. The authors are grateful to Michael Kagay of the *New York Times*, Kathleen Frankovic of CBS, David Broder of the *Washington Post*, and Larry Hugick, Jim Shriver, and Colleen McMurray of Gallup for providing materials, and to Kathleen Frankovic and Louis Harris for several helpful comments on an earlier version of the manuscript. Obviously, none of these individuals and institutions is responsible for—or necessarily agrees with—the authors' interpretations of their data or evaluations of their efforts.

Public Opinion Quarterly Volume 53:563–589 © 1989 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/89/0053-04/$2.50

Times combine.[1] Our analysis is far from exhaustive, even with respect to this very limited set of polling operations. But then, our goal is not to render a balanced summary verdict about the competence or utility of these specific organizations' efforts in the 1988 campaign. Rather, our aim is to use these organizations' experiences to illustrate some of the possibilities and pitfalls facing any polling organization in a presidential primary campaign. In a concluding section we briefly address some more general issues regarding the role of polling in the presidential nominating process.

The Primary Season Polls

The thrusts of two major primary season polling efforts are evident in Tables 1 and 2, which list the 1988 state and national campaign polls conducted by the Gallup Organization[2] and by the CBS News/*New York Times* combine from the beginning of the calendar year through the end of the primary season in early June.

The Gallup Organization conducted early and late preprimary (or precaucus) polls in each of eight key states. Most of the early polls were conducted from three to five weeks before the primary, while the

---

[1] Our choice of these three organizations was prompted in part by their public prominence among the large number of organizations engaged in polling during the 1988 presidential primary season, and in part by the ready availability of materials documenting their efforts.

[2] The phrase "Gallup Poll" is used to describe two different departments at the Gallup Organization, Inc., but only one set of polls. The Gallup corporation contracted with a consortium of television stations (and some newspapers) called CONUS to conduct political polling in 1988. The Gallup/CONUS materials given to us for review were produced under this contract. The poll interpretations were released every Monday for broadcast on Tuesday evening. Larry Hugick and Andy Kohut compiled the data and wrote the stories under their own byline.

The second use of the term "Gallup Poll" refers to the syndicated article under the byline of George Gallup, Jr., and Alec Gallup. Jim Shriver wrote these articles, which were released for use in newspapers the morning after CONUS broadcast them. Newspapers purchasing the syndicated column were free to use all, part, or none of the text released by the Gallup Organization. At large metropolitan newspapers local reporters typically rewrote the article under their own bylines, picking data and prose freely from the Gallup release.

In 1988 the Gallup Organization also provided data under contract to *Newsweek* and the Times-Mirror newspaper chain. Polls conducted under these arrangements were not referred to as "Gallup Polls," but as, for example, "*Newsweek* Polls conducted by the Gallup Organization."

In this article we use the label "Gallup" to refer to the syndicated Gallup/Shriver column and "Gallup/CONUS" to refer to the Hugick/Kohut release for the CONUS consortium. We have not analyzed the use of Gallup data by *Newsweek* or the Times-Mirror chain.

We are grateful to Jim Shriver and Larry Hugick of the Gallup Organization for explaining this organizational structure to us and for making their offices available to our inquiries.

Table 1. Gallup Polls Conducted During the 1988 Primary Campaign

Dates	Coverage	Sample Size[a] Reg	Dem	Rep	Type
1–3 Jan	Iowa	1,223	588	519	Telephone
8–10 Jan	New Hampshire	1,214	416	545	Telephone
15–17 Jan	South	1,201	370	322	Telephone
22–24 Jan	National	1,210	560	538	Telephone
29–31 Jan	Iowa	1,215	577	529	Telephone
12–14 Feb	New Hampshire	3,058	1,177	1,531	Telephone
19–21 Feb	South	1,208	426	333	Telephone
19–21 Feb	Massachusetts	1,227	639	331	Telephone
26–28 Feb	Florida	1,206	489	396	Telephone
26–28 Feb	Texas	1,207	483	496	Telephone
4–6 Mar	Florida	3,019	1,173	1,024	Telephone
4–6 Mar	Massachusetts	3,043	1,537	686	Telephone
4–6 Mar	Texas	3,078	1,302	1,139	Telephone
10–12 Mar	National	1,211	535	567	Telephone
25–27 Mar	New York	1,205	421	329	Telephone
8–10 Apr	Ohio	1,201	490	—	Telephone
15–17 Apr	New York	3,065	1,096	—	Telephone
21–23 Apr	National	1,204	584	—	Telephone
29 Apr–1 May	Ohio	3,033	1,556	—	Telephone
5–7 May	California	1,205	487	—	Telephone
13–15 May	National	1,204	608	—	Telephone
3–5 Jun	California	3,052	1,376	—	Telephone

NOTE: For final preprimary polls in New Hampshire, Florida, Massachusetts, Texas, New York, and Ohio, preliminary results were released based on first two nights' polling, followed by final results based on full sample.

[a] "Reg" refers to numbers of registered voters. "Dem" and "Rep" refer to numbers of registered party identifiers and "leaners" in national polls, and to numbers of registered voters who "plan to vote" in parties' primaries in state polls.

late polls were invariably conducted on the weekend preceding a Tuesday primary. The emphasis on preprimary polls (and especially on late preprimary polls) presumably reflects and capitalizes upon Gallup's status as the best-known "election caller" on the American political scene; scores of newspapers headlined Gallup's late results in their final day or two of preprimary coverage.

By contrast, the CBS News/*New York Times* combine conducted preprimary or precaucus polls on only five occasions: once in Iowa, twice in New Hampshire, once in the South before Super Tuesday, and

Table 2. CBS News/*New York Times* Polls Conducted During the 1988 Primary Campaign

		Sample Size[a]			
Dates	Coverage	Tot	Dem	Rep	Type
2–5 Jan	Iowa	3,126	502	432	Telephone
17–21 Jan	National	1,663	602	438	Telephone
30–31 Jan[b]	National	1,187	447	319	Telephone
25–30 Jan[c]	New Hampshire	1,784	434	513	Telephone
8 Feb	Iowa	—	1,674	1,652	Entrance
9–13 Feb[cd]	New Hampshire	5,245	1,442	1,636	Telephone
13–15 Feb[ce]	New Hampshire	—	954	1,184	Telephone
16 Feb	New Hampshire	—	1,510	1,567	Exit
17–21 Feb	National[f]	2,734	933	684	Telephone
23 Feb[c]	South Dakota	—	934	1,014	Exit
28 Feb–2 Mar	South	1,922	666	421	Telephone
8 Mar	South	—	8,259	6,681	Exit
8 Mar	Florida	—	1,494	1,384	Exit
8 Mar	Texas	—	1,647	1,276	Exit
15 Mar	Illinois	—	1,593	854	Exit
19–22 Mar	National	1,654	?	?	Telephone
5 April	Wisconsin	—	934	—	Exit
11–13 April[c]	New York	2,115	581	—	Telephone
19 April	New York	—	2,100	—	Exit
26 April	Pennsylvania	—	2,095	—	Exit
3 May[c]	Indiana	—	1,375	—	Exit
3 May	Ohio	—	1,540	—	Exit
9–12 May	National	1,382	—	—	Telephone
7 Jun	California	—	1,894	—	Exit
7 Jun	New Jersey	—	761	—	Exit

[a] "Tot" refers to total numbers of respondents; "Dem" and "Rep" refer to numbers of registered voters who "usually vote or are likely to vote this year in a primary or caucus of that party" or, in the case of exit polls, to actual primary voters.

[b] Special "Bush/Rather Survey"; *New York Times* only.

[c] CBS News only.

[d] Tracking polls conducted 9–10 Feb, 10–11 Feb, and 12–13 Feb; results released separately.

[e] Tracking polls conducted 13–14 Feb and 14–15 Feb among "likely primary voters"; results released separately.

[f] Oversampling in Super Tuesday Southern and Border states produced 575 Democratic and 375 Republican primary voters in those states.

once in New York. Most of the combine's state or regional polls (a total of 14) were exit polls (or, in the case of the Iowa caucuses, an entrance poll) intended to explain after the fact why people voted the way they did.[3] Typically, results from these polls were used by CBS News during election night coverage and by the *New York Times* in a long analytical story on the Thursday following each Tuesday's primary.

Both Gallup and the CBS News/*New York Times* combine conducted occasional national polls in addition to their state polls; in Gallup's case there were four national polls at roughly five-week intervals beginning in late January, for the combine there were four plus a special "Bush/Rather Survey" conducted in late January for the *New York Times* only.

The Louis Harris organization took a different tack, eschewing state polls entirely and conducting a total of seven national polls at roughly four-week intervals from early January through early June. The Harris polls were conducted by telephone with total sample sizes ranging from 1,018 to 1,263.[4] During the same period Harris produced a total of 45 press releases.[5] Given this disparity between the frequency of polling and the frequency of press releases, it should not be surprising that the data cited in a typical Harris release were less than ideally current. We looked at the length of time between the last day of polling and the date of the press release for 67 Harris Poll and Gallup Poll stories between 1 January and 9 June. The Harris stories referred to polls conducted, on average, 15.3 days earlier; the Gallup stories referred to polls conducted, on average, 7.4 days earlier.[6] Only 20% of the Harris stories reported data within 5 days of the last polling; the comparable figure for the Gallup Poll was 70%. In short, the Harris Poll was often reporting data that were two or more weeks old in a political world where the accuracy of surveys is highly (and negatively) correlated with the amount of time between the survey and the election (Crespi, 1988:167).

Moreover, some of the Harris Poll stories had lead paragraphs or headlines that implied the data were current when they were actually

---

[3] We exclude from Table 2 an additional series of 12 state exit polls conducted in Super Tuesday states by CBS News only; these were presumably used during election night coverage to help "call" individual primary outcomes, but we had neither the data themselves nor any other record of their use.

[4] Numbers of registered voters and breakdowns by party were not printed on the Harris press releases. However, the information is publicly available from the Harris Poll organization.

[5] The releases contain both weekly commentary and weekly polling reports. We did not differentiate between the two because the format and content of the releases was the same for both types of press release.

[6] The median length between final polling and release was 13 days for the Harris Polls and 4 days for the Gallup Polls.

Table 3. Level of Statistical Analysis in Poll Stories

	Gallup	CBS/*NYT*	Harris
Narrative only[a]	0%	0%	33%
Univariate analysis			
Frequency[b]	18	0	22
Subgroup frequency[c]	5	0	36
Bivariate analysis			
Longitudinal frequency[d]	18	8	7
Zero-order cross-tab[e]	27	72	0
Trivariate analysis			
Longitudinal cross-tab[f]	14	12	2
First-order cross-tab[g]	18	8	0
Correlation/regression[h]	0	0	0
	100%	100%	100%
N	22	25	45

[a] No tabular data included in press release.

[b] Frequency distribution of a single variable.

[c] Frequency distribution within a subgroup of the sample (e.g., preference distribution among Republicans).

[d] Frequency distributions for a single variable compared at two or more points in time.

[e] Cross-tabulation with a single explanatory variable (e.g., vote preference among liberals, moderates, and conservatives).

[f] "Zero-order" cross-tabulations compared for two or more points in time.

[g] Cross-tabulations with control variables other than time.

[h] Correlation or regression coefficients.

several weeks old. On 3 March, five days before Super Tuesday, Harris released a report headlined: "Negative Vote Pivotal in Super Tuesday Balloting." The headline explicitly indicated that the poll had relevance for the upcoming Super Tuesday primaries. But the survey on which the report was based was a national survey conducted immediately after the New Hampshire primary. These data were appropriate for measuring the national "bounce" Dukakis and Bush received from their victories in the Granite State, and were used that way in two previous Harris releases. But they told us little about Super Tuesday—certainly less than was already available from weekly tracking of opinion in state polls in the Super Tuesday states.

The lengthy period between polling and use of the data in the case of the Harris Poll is apparently due to a scarcity of data rather than to any special sophistication of analysis. Table 3 shows the level of statistical sophistication employed in reports of results from the Harris, Gallup, and CBS News/*New York Times* polls. The Harris Poll syndicated column relied on narrative with no tabular presentation of data at all in one out of every three press releases. By contrast, the Gallup Poll always provided marginals, and presented at least one cross-tabulation (with a demographic variable, party, time, or some other explanatory variable) in half its releases. The most common analysis of the CBS News/*New York Times* data compared the cell frequencies of cross-tabulations of an explanatory variable specific to a particular survey (i.e., not demographics, party, or time, which are found in all combine surveys). None of the poll stories of the three polling agencies in our analysis used descriptive statistics such as means or medians, not to mention correlation or regression coefficients, to describe their data.

The Early Line: Measuring Favorability

Campaigns for the presidency now begin a year or more before the first caucuses and primaries. The pollsters become involved as best they can, diligently tapping public opinion long before most of the public has given the campaign anything more than minimal thought. The chief visible product of this preseason polling is a pecking order of first choice support for each party's nomination. Who is the front runner? Who are the major challengers? But these early readings reflect name recognition more than potential support in a real campaign setting. A large fraction of the electorate typically reports being undecided, and whatever preferences are expressed are notoriously volatile. The addition or subtraction of a well-known fence sitter (Ted Kennedy, Mario Cuomo) can scramble the results, while relative unknowns may later emerge from nowhere ("an asterisk in the polls") to win the nomination.

These limitations have prompted many pollsters to supplement reports of first-choice support with additional information intended to provide a more valid fix on each candidate's latent electoral potential. In 1988 Gallup tried a variety of approaches: a question inviting each respondent to designate a second choice for the nomination from the full list of candidates, a question probing for "strong opposition" to any of the candidates, and (most frequently) a question asking whether respondents supported their first choice "strongly or only moderately."

The CBS News/*New York Times* polls relied throughout the campaign on a separate measure of favorable and unfavorable impressions of each candidate. This format has the disadvantage of requiring a separate question for each candidate. But it has the more-than-

Table 4. Favorability Ratings and First Choices Among Democrats in CBS News/*New York Times* Primary Season Polls

	January		February		March	
	Favorable	Support	Favorable	Support	Favorable	Support
Dukakis	74	6%	79	21%	76	29%
Jackson	54	17	50	13	57	22
Gore	61	4	39	8	51	14
Gephardt	57	4	63	12	34	8
Simon	73	9	44	6	39	6
Hart	42	23	24	10	n.a.	n.a.
Babbitt	54	2	n.a.	1	n.a.	n.a.
Other, none, DK		35		30		22
		100%		100%		100%

SAMPLE: Registered voters who "usually vote or are likely to vote this year" in a primary or caucus of the candidate's party.

NOTES: "Favorable" entries are percentages of favorable ratings among all favorable and unfavorable ratings for each candidate; "undecided" and "haven't heard enough to have an opinion" responses are ignored in calculating these percentages. "Support" entries are percentages of first choice support.

compensating advantage of providing candidate-specific readings relatively independent of the composition of the field of candidates as a whole. In addition, the ratings measure how much the public knows about each candidate (by comparing the number of favorable plus unfavorable responses to the number who are undecided or who "haven't heard enough to have an opinion"), as well as the balance of favorable and unfavorable responses among those who do have an opinion.

The potential value of such a format, especially in the early stages of a crowded and relatively undefined race, is illustrated in Table 4, which compares the favorability ratings and first choice support of the Democratic candidates in 1988 CBS News/*New York Times* polls. The candidates are listed in order of their standings in the March preference poll. But there is virtually no correspondence between preferences in March and January; Dukakis went from fourth place in January to first place in March, Hart went from first place to no place, and so on. (For connoisseurs of statistical wizardry, the proportion of intercandidate variance in March preference standings accounted for by January preference standings is less than 1%—.007.) By contrast, favorability ratings were both more stable over time (except for candidates, like Hart in February and Gephardt and Simon in March, who had all but dropped out of the race) and more strongly related to the candidates'

Table 5. Balance of Favorable and Unfavorable Ratings of Candidates in Early CBS News/*New York Times* National Polls

	Oct 1987	Nov 1987	Dec 1987	Jan 1988	Average
Bush	84	83	n.a.	80/76	81
Dole	86	81	n.a.	83/67	79
Kemp	62	55	n.a.	69	62
Robertson	16	21	n.a.	23	20
Dukakis	70	64	86	74	73
Simon	70	65	66	73	68
Gephardt	67	47	58	57	57
Gore	68	45	50	61	56
Jackson	38	43	46	54	45
Hart	n.a.	n.a.	59	42/34	45
Babbitt	40	25	33	54	38

SAMPLE: Registered voters who "usually vote or are likely to vote this year" in a primary or caucus of the candidate's party.

NOTES: Entries are percentages of favorable ratings among all favorable and unfavorable ratings for each candidate; "undecided" and "haven't heard enough to have an opinion" responses are ignored in calculating these percentages. Second January entries for Bush, Dole, and Hart are from the special *New York Times* "Bush/Rather Survey" conducted 30–31 January 1988.

eventual preference standings in March. As early as January, a ranking of Democratic candidates by the proportion of favorable impressions would have had Dukakis first, Hart last (in spite of his lead in the January preference poll), and Gephardt and Gore in the middle of the pack.[7]

Table 5 shows the balance of favorable and unfavorable ratings in CBS News/*New York Times* polls for Republican and Democratic candidates from October 1987 through January 1988. Even in this earliest phase of the public campaign the favorability ratings were reasonably good predictors of the candidates' performances in the spring. On the Republican side Bush and Dole were ahead but fading, with Dole fading faster than Bush. On the Democratic side Dukakis was first, with Babbitt last, Hart low and dropping steadily, and Jackson gaining steadily. (We look more closely at the Jackson case later in this review.)

---

[7] The favorability ratings shown in Tables 4 and 5 were computed from percentages reported in CBS News press releases; calculations based on raw data would produce slightly different ratings by avoiding rounding error.

The 1988 primary season had plenty of surprises in store, even for a careful follower of these early favorability results—but not nearly as many as for someone focusing on the more highly publicized "horse race" numbers in the early preference polls. If one aim of the pollsters and poll interpreters is to reduce our quadrennial sense of surprise at the events of the presidential nominating process, greater emphasis upon (and more thorough analysis of) early favorability data might help to achieve that aim.

The Accuracy of Preprimary Polls

Political pollsters like to say that a poll result is not a prediction, but merely a "snapshot" of voter preferences at a specific time some days or weeks before an election. This becoming modesty notwithstanding, public pollsters' salaries are paid directly by media executives (and indirectly by readers and listeners) whose interest in election polls rests squarely on the belief that polls predict what will happen on election day.

Then too, pollsters have not hesitated to boast when their "snapshots" of public opinion did turn out to be accurate predictors of actual election results. A *New York Times* article about the accuracy of the polls began:

> After some embarrassing findings in the Presidential primary season, newspapers came about as close as polling specialists believe is possible to forecasting the results of Tuesday's election.
>
> That success has not silenced the debate over how much polling is too much; news organizations conducted nearly three times as many national surveys this year as in the 1984 general election. . . .
>
> "The pollsters certainly did a strong job this time," said Everett Ladd, director of the Roper Center for Public Opinion Research at the University of Connecticut. (Rosenthal, 1988:17)

Thus, the pollsters themselves suggest that the correspondence between poll results and actual election outcomes is an appropriate criterion for evaluating election season polls.

As we have already seen, neither Harris nor the CBS News/*New York Times* combine was heavily involved in preprimary polling in specific states during the 1988 campaign. That role was left to the Gallup Organization, which did conduct preprimary (or precaucus) polls in eight key states: Iowa, New Hampshire, Florida, Massachusetts, Texas, New York, Ohio, and California. How accurate were these preprimary polls as forecasts of the primary outcomes?

The answer to that simple-sounding question is complicated by Gallup's production of three distinct sets of poll results in each state. Early polls conducted from three to five weeks before each primary produced one set of results, late polls conducted over the weekend preceding each primary produced a second set of results, and subsamples of "most likely voters" drawn from the late polls in each state produced a third set of results.[8] In Table 6 we summarize the discrepancies between each of these three distinct sets of poll results and the actual primary results.

The average absolute (unsigned) error in predicting Bush's and Dukakis's vote percentages from the full sample of respondents "intending to vote" in late Gallup polls in the 12 primaries and caucuses represented in Table 6 was 5.9 percentage points. Discrepancies of this magnitude are somewhat larger than would be expected simply on the basis of sampling error; nevertheless, they seem quite respectable given "the practical difficulties of conducting any survey of public opinion" (a standard industry disclaimer) plus the peculiar complications and volatility of primary season polling.

Among respondents judged "most likely to vote" the corresponding average absolute error for the same 12 state polls was even smaller, about 4.3 percentage points. It appears from this result that Gallup's efforts to identify likely voters were more than sufficiently successful to outweigh their costs in terms of reduced sample size.[9] The corresponding errors for the earlier preprimary polls were somewhat larger, averaging about 7.6 percentage points. Even these results seem quite respectable given the amount of time between the polls and the primaries (from three to five weeks) and the markedly smaller sample sizes in Gallup's early preprimary polls.[10]

In addition to showing the average *absolute* errors for Gallup's preprimary polls, Table 6 also shows the average *signed* error for each candidate. From these results it appears that Gallup systematically overestimated support for Robert Dole (by averages of 2.0, 3.4, and 2.8 percentage points in the three sets of poll results) while underestimating support for Pat Robertson (by from 4.5 to 5.6 percentage points) and Jesse Jackson (by from 4.3 to 6.4 percentage points).

Both Robertson and Jackson drew heavily from relatively poor rural

---

[8] Gallup made no attempt to identify "most likely voters" in the early preprimary polls. The organization's own analyses of the late polls attached varying degrees of emphasis to the results based on all respondents "planning to vote" and those based on respondents "most likely to vote."

[9] The average errors shown in Table 6 are based on average sample sizes of 1,069 for respondents "planning" to vote and 540 for respondents "most likely" to vote; Gallup's methods for selecting "most likely" voters (never clearly documented in the press releases) thus involved cutting effective sample sizes nearly in half.

[10] The average sample sizes were 1,069 in the 12 late polls and 484 in the 12 early polls included in Table 6. In theory, that difference alone would make the expected average errors in the early polls about 50% greater than those in the later polls.

Table 6. Accuracy of Gallup's Preprimary Polls

	Average Absolute Error	Average Signed Error
	Early Preprimary Polls: Respondents "Planning" to Vote	
Bush (*N* = 5)[a]	7.4	−3.1
Dole (*N* = 5)	4.8	+2.8
Robertson (*N* = 5)	5.8	−5.6
Dukakis (*N* = 8)[b]	7.6	−4.7
Jackson (*N* = 8)	6.4	−6.4
Bush/Dukakis (*N* = 12)[c]	7.6	−4.9
	Late Preprimary Polls: Respondents "Planning" to Vote	
Bush (*N* = 5)	5.9	−2.5
Dole (*N* = 5)	3.0	+2.0
Robertson (*N* = 5)	5.1	−4.7
Dukakis (*N* = 8)	5.5	−0.8
Jackson (*N* = 8)	5.5	−5.5
Bush/Dukakis (*N* = 12)	5.9	−1.8
	Late Preprimary Polls: Respondents "Most Likely" to Vote	
Bush (*N* = 5)	3.3	−2.7
Dole (*N* = 5)	4.7	+3.4
Robertson (*N* = 5)	5.1	−4.5
Dukakis (*N* = 7)[d]	5.3	−1.1
Jackson (*N* = 7)	4.3	−4.3
Bush/Dukakis (*N* = 12)	4.3	−1.8

[a] States in which Gallup reported Republican preprimary poll results included Iowa, New Hampshire, Florida, Massachusetts, and Texas.

[b] States in which Gallup reported Democratic preprimary poll results included Iowa, New Hampshire, Florida, Massachusetts, Texas, New York, Ohio, and California.

[c] To facilitate comparison across types of polls, California primary poll results are excluded from Bush/Dukakis summary calculations.

[d] No subsample results for "most likely" voters were reported from Gallup's late preprimary poll in California.

constituencies in which survey sampling is especially difficult.[11] Thus, some of Gallup's undercounting of these candidates' support is presumably due to the fact that their supporters simply failed to turn up in the Gallup sample in proportion to their numbers. Robertson and Jackson also both claimed special strength among voters who were not part of the traditional electorate; such voters are presumably most likely to be written off as unlikely to vote.

The undercounting of Robertson and Jackson's support has some methodological implications beyond the findings of these particular surveys. Widely used survey techniques, such as random digit dialing, cluster sampling, or reweighting to reflect a "probable electorate" can misrepresent the support of candidates whose constituencies are especially sensitive to any bias created by these techniques. Even careful demographic weighting is unlikely to recreate the opinions that are missed in an unrepresentative sample.[12] In the 1988 primary campaign, support for Robertson and Jackson was probably that sort of opinion.

Tracking the Race: The Case of New Hampshire

For the pollsters, the most embarrassing episode of the 1988 primary season was George Bush's remarkable comeback in the final days of the New Hampshire primary campaign. Bush had led Robert Dole by margins approaching 2 to 1 in the month before the primary. But his lead evaporated after Dole's victory (and Bush's third-place finish) in the much-publicized Iowa caucuses. Most late polls in New Hampshire had Dole even with Bush or pulling away. None forecast a clear Bush victory. But Bush did win clearly, by more than nine percentage points. At least one pollster was fired from a national account for having had things so wrong. "I've studied data from thousands of elections," another pollster told Paul Taylor (1988:15) of the *Washington Post*. "This is the most unique finish I've ever seen."

Both the Gallup Organization and CBS News conducted tracking polls during the final week of the New Hampshire primary campaign. Thus, the campaign's unexpected finish provides an interesting opportunity to assess each organization's results—and each organization's

[11] For example, 29.2% of households with incomes below $5,000 had no telephone in 1985–86; 15.6% (19.9% in the South) of black households had no telephone. Obviously, telephone surveys omit these people. See Thornberry and Massey, 1988:34–35.

[12] Others have noted that "a poststratification realignment by age, sex, and education will not align the telephone population very well by race and region." These were, of course, two important variables in Jackson's strength. See Massey and Botman, 1988: 160. However, also see Mitofsky, 1981.

interpretation of those results—in a treacherously volatile and politically crucial setting.

The results of the CBS News polls in New Hampshire, together with excerpts from the polling operation's interpretations of those results, are shown in Table 7. CBS News conducted a total of five separate polls in the week between the Iowa caucuses and the New Hampshire primary, the last conducted on the Sunday and Monday before the primary (14 and 15 February) and released for use on the network's Tuesday morning news show. The cumulative Republican sample size for these five polls was 1,964.

The Gallup organization polled from Friday (12 February) through Sunday (14 February). Preliminary results based on 993 Republicans interviewed Friday and Saturday were released for publication in newspapers Monday morning, and cumulative results based on 1,531 Republicans interviewed through Sunday were released as an update for publication on Tuesday morning. The daily results (included in the second release), plus excerpts from the Gallup/CONUS interpretation included with each release, are shown in Table 8.

One crucial difference between the CBS News and Gallup polls is evident from a comparison of the raw percentages in Tables 7 and 8. CBS News never had Bush trailing Dole by more than three percentage points throughout the week before the primary, and its final poll on Sunday and Monday had Bush pulling out to a four-percentage-point lead. Gallup had Dole ahead by eight to nine percentage points (among "likely voters") all week long, with little sign of movement from day to day. Which organization "had it right" early in the week is, of course, impossible to say. But CBS News was clearly much closer to the mark later in the week, where the inevitable benchmark of success is an accurate forecast of the actual primary outcome.

Ironically, given what turned out to be better data, CBS News produced "in house" analyses of those data that were noticeably more circumspect than the corresponding "in house" analyses of the Gallup data.[13] The main thrust of the CBS News analysis throughout the week, which we have attempted to convey with brief excerpts in Table 7, was that the race was "too close to call." When Bush emerged with a four-point lead in the final round of polling, the CBS News analysis emphasized the slimness of that lead and the fact that it was based on seemingly "soft" support. By contrast, the first post-Iowa Gallup/CONUS report announced a "statisically significant" Dole lead, and subsequent reports emphasized the stability of that lead from day to day: the main impression conveyed by these reports was one of a

---

[13] Here we refer to the Gallup/CONUS analyses produced by Larry Hugick and Andy Kohut which were provided to us for evaluation.

Table 7. Republican Horse Race in New Hampshire in CBS News Tracking Polls

Dates	Bush	Dole	*N*	Commentary
25–30 Jan	42%	20%	513	"Bush has a two to one lead. . . . Most Republicans say they could change their minds between the interview and election day, but over half of Bush's supporters say their minds are made up."
9–10 Feb	35%	27%	543	"In the two days after the [Iowa] caucuses, Bush's lead has dropped to only eight points. . . . In the current poll, Bush's lead narrowed by a significant margin between Tuesday and Wednesday."
10–11 Feb	29%	32%	499	"George Bush is no longer the frontrunner in New Hampshire. Interviews . . . now show a race between Bob Dole and George Bush that is too close to call. . . . CBS News' continuing New Hampshire polling clearly indicates Bush's support declining and Dole's rising."
12–13 Feb	30%	31%	594	"still locked in a tight race . . . virtually no change."
13–14 Feb	32%	31%	588	"still locked in a head-to-head confrontation . . . support for each frontrunning candidate is about equal."
14–15 Feb	34%	30%	596	"still in a close race . . . Bush's slim lead is within the margin of error and comes entirely from those . . . who say they will 'probably vote' in the Republican primary. Among those who say they will 'definitely vote' in the primary, Bush gets 31% and Dole gets 31%."
16 Feb	37.6%	28.4%	157,625	

Table 8. Republican Horse Race in New Hampshire in Gallup Polls

Dates	Bush	Dole	N	Commentary
8–10 Jan	38%	23%	545	"Bush outdistances his closest GOP rival, Bob Dole, by almost a two-to-one margin. . . . Bush support is both wide and deep. . . . The 'Iowa effect' on the race, of course, cannot be factored into the results."
12–13 Feb	28% 28%	33% 36%	993 589[a]	"The wide lead Vice President Bush held over Senator Bob Dole among New Hampshire primary voters one month ago is gone. . . . Dole now leads over Bush (36% to 28%) among most likely Republican primary voters. The lead is statistically significant."
12–14 Feb	27%	35%	902[a]	"Sunday's interviewing finds no significant change in candidate preferences among New Hampshire voters most likely to participate in Tuesday's Republican presidential primary. Based on all likely GOP primary voters interviewed by Gallup Friday through Sunday, Bob Dole leads George Bush by a margin of 35% to 27%. . . . The daily trend in candidate support provides no evidence of any late movement in the GOP race."
12 Feb	27%	35%	321[a]	
13 Feb	29%	38%	265[a]	
14 Feb	25%	33%	315[a]	
16 Feb	37.6%	28.4%	157,625	

[a] Likely voters.

sudden and seemingly irreversible shift to Dole following his Iowa triumph.

The Gallup/CONUS analysis of "strong" and "not strong" support did reveal some of the vulnerability of Dole's lead:

> Although Dole continues to hold a significant lead over Bush, Dole's support remains less firm than his rival's. The majority of likely voters who support Bush strongly back his candidacy compared with fewer than half of Dole's supporters. To maintain his advantage and win on Tuesday, Dole must hold on to his less committed supporters.

Of course, that is what Dole failed, in the event, to do. But neither the Gallup organization's own poll results nor the (Gallup/CONUS) interpretation provided grounds for anticipating that failure.

So far we have examined only an "in house" analysis of the Gallup Organization's New Hampshire data. But working journalists' use of those data represents another, and perhaps more consequential, product of the polling enterprise. We focus here, more or less at random, on a front-page story by Jerry Roberts in the 15 February edition of the *San Francisco Chronicle*, a story headlined "Gallup Poll Says Dole Leads in New Hampshire."[14] "One day before the nationally watched New Hampshire primary," Roberts wrote,

> Dole leads Bush 36 percent to 28 percent among Republicans considered most likely to vote, according to the Gallup Organization survey. . . . The new survey mirrors two other New Hampshire polls released yesterday, by the Boston Globe and the Washington Post–ABC. . . . Both polls also showed Dole leading Bush, but by small margins that were statistically insignificant. Dole's lead in the Gallup Poll, however, is statistically significant.

Roberts's story illustrates several important features of the way working journalists often use poll results. Taken together, these practices tended in this instance (as they do more generally) to exaggerate both the certainty and the political significance of the poll results on which the story was based:

1. Roberts followed the Gallup/CONUS press release's lead in focusing on Dole's more dramatic eight-point lead among "likely voters"; indeed, Roberts's story made only passing reference to Dole's smaller five-point lead in the larger sample. The accompanying table included both margins, but misreported the former as being based on registered voters and the latter as being based on those planning to vote in the primary (thus exaggerating the apparent change in Bush's fortunes since January).
2. Roberts's story read as though Dole's lead existed on the day of publication—"one day before" the primary. At that point the data on which the story was based were two to three days old. Roberts buttressed the Gallup results by citing congruent data from other surveys—surveys also released on 14 February but conducted earlier.
3. Roberts referred to the Gallup results as being "statistically significant," but he included no discussion of what that meant,

---

[14] Roberts's story happened to be near to hand. Our assumption, though we have no systematic evidence to support it, is that similar examples could have been drawn from almost any major newspaper's coverage of almost any important primary.

or of the possibility that the advertised "margins of error of 3 percent to 5 percent" might understate the actual uncertainty of the results in a volatile primary setting.

4. A cautionary sentence indicating that "Dole's support is not as firm as that of Bush" appeared as the sixteenth paragraph of Roberts's story, far below the description of the "breathtaking suddenness" of "Bush's nose dive." As is often the case in journalism, clarification and qualification were not allowed to intrude significantly upon a clear, compelling storyline.

The combined effect of these several seemingly insignificant journalistic decisions was to obscure for the general reader of Roberts's story the (already deemphasized) elements of uncertainty in the original Gallup/CONUS analysis, and thus to increase rather than decrease the potential for surprise in the next day's primary result.[15]

Breaking New Ground: The Case of Jesse Jackson

The candidacy of Jesse Jackson, another surprise for political pundits in 1988, created special problems for the media and pollsters. There was simply no precedent, no baseline, no history for comparison. Jackson was the first black candidate with a chance to gain a major political party's nomination for president.

Until early March the press seemed unaware of Jackson's historic role. But a surprisingly strong showing in several Super Tuesday primaries, followed by another strong showing in the Illinois primary and then a startling victory in the Michigan caucuses, produced a second phase in the press and pollsters' treatment of Jackson. He was recognized as a serious contender for the nomination. After Jackson lost New York his nomination once again became improbable, and the press and pollsters began a third phase of coverage centering primarily on speculation about the vice-presidential nomination and about Jackson's future role in the Democratic party.

The three phases of Jackson's candidacy are reflected in the questions asked by the polling agencies between January and June, some of which are shown in Table 9. In contrast to television news stories about Jackson's 1984 bid for the presidency (see Broh, 1987), poll questions during the first phase in 1988 consistently included Jackson as part of the horse race. Throughout January and February, Gallup, Harris, and the CBS News/*New York Times* combine had Jesse Jackson as a response to their trial heats and exit polls. Standard survey

---

[15] While the evidence here is clearly anecdotal, the pattern of reporting is consistent with content analyses from other presidential elections. See Broh, 1980.

Table 9. Three Phases of Jesse Jackson Poll Questions

Phase 1: Horse race and candidate assessment

Gallup: Regardless of which Democratic presidential candidate you happen to prefer, which one do you think cares most about people like yourself?

Harris: Let me read you a list of some people who have been mentioned as possible Democratic candidates for president in 1988. Who on that list do you feel you are not familiar with?

CBS/NYT: Who do you want the Democrats to nominate for President in 1988—Michael Dukakis, Dick Gephardt, Albert Gore, Gary Hart, Jesse Jackson, or Paul Simon? Is your mind made up, or is it still too early to say for sure?

Phase 2: "Assessing a contender"

Gallup: Is your opinion of Jesse Jackson very favorable, mostly favorable, mostly unfavorable, or very unfavorable?

Harris: Let me read you a list of some people who have been mentioned as possible Democratic candidates for president in 1988. Which people on that list do you feel you could not vote for if nominated for president in 1988?

Harris: Agree/Disagree: Because of his radical tendencies, it's dangerous for a person like Jesse Jackson to get too much power.

Harris: Agree/Disagree: The country isn't ready yet for a black to be president.

Harris: Agree/Disagree: There are real questions about how honest [Jesse Jackson] is on money matters and where his money is coming from.

Harris: Agree/Disagree: He makes more sense when he debates than most of the other candidates for president in both parties.

CBS/NYT: Regardless of which candidate you support for the nomination, which of these Democratic candidates do you think would have the *best chance of winning* the election in November if he were nominated—Michael Dukakis, Dick Gephardt, Albert Gore, Jesse Jackson, or Paul Simon?

Phase 3: Vice-presidency and future role

Gallup: If the November election were being held today and the Republican candidates were George Bush for president and Jack Kemp for vice president, while the Democratic candidates were Michael Dukakis for president and Jesse Jackson for vice president, who would you like to see win?

Harris: Now, on the Democratic side, most observers are convinced that Governor Michael Dukakis of Massachusetts will win the Democratic nomination for president this July. Of the following people, which one would you prefer as your first choice for the Democratic candidate for Vice President to run with Governor Dukakis?

CBS/NYT: Do you think Jesse Jackson has too much influence on the Democratic party, not enough influence, or about the right amount of influence?

questions about handling the nation's problems, chances of winning, and caring about the voters also included Jackson in the wording.

But as in 1984, Jackson was largely unreported in the first phase of the 1988 primaries. For example, Jackson figured predominantly in only one press release of any agency—a Louis Harris (1988) essay (with no current data) about the "wild cards" in the election.[16] Similarly, in the CBS News releases about Iowa and New Hampshire, Jackson's showing was either ignored or relegated to the last paragraph.

Super Tuesday changed the 1988 Democratic contest. Jackson won enough primaries for the press and pollsters to consider him a contender for the nomination.[17] Meg Greenfield (1988) editorialized in *Newsweek* that anything less was racism, supporting David Broder's (1988) earlier editorial that the press had to change its reporting of Jackson. *Time* (Shapiro, 1988:21) and Dan Rather[18] referred to Jackson during this phase as the "frontrunner" for the Democratic nomination. Pollsters responded.

Phase two of the Jackson campaign is best summarized as "assessing a contender." Once Jackson was perceived as having a chance to win the nomination, however unlikely, he received the same rough treatment that serious challengers receive. From the pollsters' perspective, this meant measuring the positives and the negatives—the image makers' benchmark for success—and finding how many voters supported him—the politicians' benchmark for success. Table 9 displays typical question wording used by Gallup, Harris, and CBS News/*New York Times* in assessing the image of Jackson. Some of the questions were part of an index containing both positive and negative attributes; others were neutral in tone. Some were standard favorability and perceived voter strength questions. But all had the common characteristic of taking Jackson's image with voters very seriously.

A major use of exit polls during this period was to monitor Jackson's share of the white primary vote to evaluate his claim for support from many races—the so-called Rainbow Coalition. After receiving 7% of the white vote in CBS News/*New York Times* exit polls in the South on Super Tuesday and again in Illinois, Jackson received 23% of the white primary vote in the Wisconsin exit poll. A breakthrough? The *New York Times*'s analysis emphasized not the dramatic gain but the glass still three-quarters empty. In a 6 April article headlined "For Jackson, Good Will, But Not Enough Votes," R. W. Apple, Jr., announced that "the dominant finding" of the Wisconsin exit poll was that "the positive feelings of many" white voters "did not translate into votes for Mr. Jackson." Apple went on to conclude that "Mr. Jackson's failure to win here seems to suggest that he has little chance of winning in the big states still to be contested." The extent to which Apple's prophecy may have been self-fulfilling is impossible to establish, but it clearly was fulfilled: Jackson went on to average only 15% of the white vote (and almost 95% of the black vote) in the five subsequent CBS News and CBS News/*New York Times* exit polls.

---

[16] The general theme of this release was that Jesse Jackson, Gary Hart, and Pat Robertson could each have become "spoilers" within their respective political parties. While this early discussion recognized Jackson's importance in 1988, it did not treat him as a contender for the nomination. In general, the Harris organization looked at Jackson's image in more depth than either the Gallup organization or the CBS News/*New York Times* combine.

[17] Jackson's victories, along with the victories of Michael Dukakis, in the Super Tuesday primaries provided an ironic twist to the nomination strategy devised by several Southern party leaders. The numerous southern states that agreed to hold their primary on the same day were supposed to form an attractive political arena for a conservative Democrat. Instead, the major beneficiaries of the day were the most liberal candidates. The Southern governors who devised the plan had forgotten that a large portion of Southern Democrats are black and that another large portion are more sensitive to national trends than to political ideology.

[18] CBS Evening News with Dan Rather, 30 March 1988, 5:33 P.M.

Phase three was the demise of the Jackson campaign and subsequent events leading up to the nominating convention. Numerous questions, experiments, and analyses about the vice-presidency took place during this phase. The end of the primary season left open the role of Jesse Jackson in the Democratic convention or in a Michael Dukakis campaign or as the vice-presidential nominee. Pollsters used the tools of their trade to ponder these issues.

A late-April Gallup Poll focusing on the upcoming general election horse race provided the Dukakis campaign with ample reason to scratch Jackson as a possible vice-presidential nominee. In a trial heat Bush won over Dukakis 45% to 43%. But when Jesse Jackson was added to the Democratic ticket against Bush and Jack Kemp as the Republican vice-presidential nominee, the percentage difference was greater: Bush-Kemp beat Dukakis-Jackson 50% to 40%.

The Jackson "damage" to a Democratic ticket was a major news item that day. George Gallup, Jr., and Alec Gallup released their syndicated column with the headline "Dukakis Runs Neck and Neck with Bush but Falls Behind with Jackson as V.P." The *San Jose Mercury News*, a respected northern California daily, ran a story about the poll on page 1A with the following lead: "Jesse Jackson's presence in the vice presidential slot on the Democratic ticket would severely damage Massachusetts Governor Michael Dukakis' chances of winning the White House, according to a new Gallup Poll" (Hamilton, 1988). The wire services carried these results the day after the Gallup news release.

Jackson caused his team to lose the trial heat; that's bad press. But it is also bad experimental design. Gallup's experiment had two simultaneous manipulations. Thus, a second plausible interpretation of the result is that Jack Kemp was a popular choice for vice-president, creating increased support for the Republican ticket. The survey design simply does not allow us to distinguish change attributable to Jackson on the Democratic ticket from change attributable to Kemp on the Republican ticket.[19] The interpretation placed on the data by Gallup (and further emphasized in the *Mercury News* story) was the least flattering one from Jackson's perspective—and one that tended to obscure the traditional journalistic concern, Who is ahead?

The CBS News/*New York Times* poll had a better research design for testing vice-presidential candidates, and it demonstrated the weakness of a Democratic ticket with Jackson as vice-presidential nominee more convincingly than did Gallup. A trial heat question pitted Republicans George Bush and Governor Jim Thompson of Illinois against (1) Dukakis and Jackson, (2) Dukakis and Nunn, and (3) Dukakis and Glenn. Like Gallup's, this experiment had no control group for comparison, but, unlike Gallup's, the design did allow comparison of several concrete alternatives on the Democratic side.[20] The results suggested that the Dukakis-Jackson ticket would fare worst against the Bush-Thompson combination. That is, Dukakis-Jackson won by only 3% while Dukakis-Nunn, Dukakis-Glenn, or Dukakis alone won by 12%, 15%, and 10% respectively; the Dukakis-Jackson ticket would apparently have produced the closest race.[21]

Despite having a more convincing experimental design, the *New York Times* gave less prominence to the results of its vice-presidential experiment. Jackson's apparent effect on the trial heat results received cursory treatment in the last paragraph of an article emphasizing the Dukakis lead:

---

[19] This methodology and question wording were not unique to the Gallup Poll. On 12 June 1988, The *Star-Ledger*/Eagleton Poll released a New Jersey poll with the headline "Dukakis Begins '88 Presidential Campaign with 14 Point Edge; Jackson Hurts Ticket in Jersey." The data showed 50% to 36%, Dukakis over Bush and 44% to 44%, Dukakis-Jackson tying Bush-Dole. Perhaps Bob Dole would have been a popular vice-presidential selection in the Garden State.

[20] Ideally, one would cross-tabulate each experimental condition with a control group to determine the impact of the experimental condition. In the absence of a control group, one can only compare one experimental condition with every other experimental condition. A question pitting Bush-Thompson against Dukakis and "some other qualified candidate for vice president" might have provided a better control for the experiment. Then the distribution of support with each of the named alternatives in the Democratic vice-presidential spot could have been compared with the distribution of support with the unspecified alternative.

[21] This result was by no means universal. A 30–31 March Roper poll showed that Dukakis-Jackson would win over Bush-Thompson 47% to 42%. With Albert Gore as the vice-presidential candidate or with no vice-presidential candidate mentioned, Dukakis and Bush were locked in a dead heat 45% to 44% and 43% to 44%, respectively. This poll received very little national coverage. A month later a Yankelovich poll showed that Dukakis-Jackson would win by the same margin, 47% to 42%. However, Dukakis alone or with Albert Gore won by larger margins, 50% to 39% and 49% to 39%, respectively. The *New York Times* ran an AP story about the poll with the daily election coverage: "Poll Gives Dukakis the Edge with Either Jackson or Gore," *New York Times*, 25 April 1988, B6.

> The survey found that with Mr. Jackson on the ticket, Mr. Dukakis's 10 point lead was cut to three points: Dukakis-Jackson got 45 percent; Bush-Thompson got 42 percent. On the other hand, Senator John Glenn of Ohio and to a lesser extent, Senator Sam Nunn of Georgia, seemed to help Mr. Dukakis. The Dukakis-Nunn ticket defeated Bush-Thompson by 46 percent to 34 percent. The Dukakis-Glenn ticket beat Bush-Thompson by 49 percent to 34 percent. (Dionne, 1988)

While the Gallup and CBS News/*New York Times* polls gave Jackson's vice-presidential bid little encouragement, the Harris Poll did provide some support. In early May, the Harris Poll included a question about future voting behavior "if a black were nominated for vice president on the Democratic ticket." Voters who said they could not support a Democratic ticket with a black vice-president were predominantly Republican, while 40% of the registered black voters said they would be more likely to show up at the polls in November if a black vice-presidential candidate were on the ticket. Harris reasoned that the added black Democrats would offset any negative reaction against a black vice-president since those most strongly opposed would probably vote Republican anyway. The same Harris Poll also showed that Jesse Jackson was the plurality choice for vice-president among registered Democrats.[22] He received 23% to John Glenn's 21% (and eventual nominee Lloyd Bentson's 6%).

Summary and Discussion

We began by asking how—and how well—pollsters dealt with the political complexity of the 1988 primary season. Perhaps not surprisingly, the complexity of the primary season itself makes for complex answers to such simple questions. Different pollsters did different things—and did different things well.

Gallup, Harris, and the CBS News/*New York Times* combine all conducted periodic national surveys during the primary season. These national polls varied in frequency and quality, as well as in the timeliness and sophistication with which they were publicized and interpreted. In general, they seem to have been most useful for establishing baselines early in the campaign season, as with the CBS News/*New York Times* combine's early readings on the favorability of the various candidates' images summarized in Tables 4 and 5. However, we were surprised by the relative infrequency of any sophisticated comparisons of data from these surveys over time, among regions, or with parallel data from individual primary states.

---

[22] The CBS/*New York Times* poll after primary season, between 20 June and 5 July, showed that Jesse Jackson was also the plurality choice among Democratic convention delegates. The *New York Times* ran a story in the politics section on 6 July 1988 with the headline "Jackson Runs 1st For the No. 2 Spot." Eighteen percent of the delegates answered "Jackson" to the question, "If Dukakis is nominated, whom do you think he should pick to run for Vice President?"

However, the *Times* poll also gave Michael Dukakis reason not to select Jackson. The percentage of delegates who said Dukakis would have an edge in their state fell from 67% to 33% when Jackson was added to the ticket; the percentage who said Bush would have an edge climbed from 9% to 37% when Jackson was added.

The Gallup Organization supplemented its national polling with a series of preprimary polls in eight key states. Those polls predicted the actual vote percentages for the leading candidates with sufficient accuracy to give both prospective primary voters and the national audience a reasonable picture of the strategic situation in most of the states where Gallup polled—but with sufficient error to leave considerable room for surprise in the closer, more volatile primaries.[23] Daily tracking polls seem especially likely to produce either boredom (if the result is clearcut) or confusion (if it is not), as Gallup—and Bob Dole—discovered in New Hampshire.[24]

The CBS News/*New York Times* combine did extensive exit polling among primary voters in key states. Since the marginals in an exit poll are never newsworthy, this approach had the considerable advantage of focusing resources and attention on the tasks of explaining and interpreting each week's results.[25] Explanation and interpretation are especially crucial during the primary season, in part because of the sheer complexity and unfamiliarity of the process, but also because each week's results can have a significant impact not only upon the public at large but also upon prospective voters in subsequent weeks' primaries.[26]

---

[23] Recall from Table 6 that Gallup's average absolute errors for Bush and Dukakis were 7.6 percentage points in early polls (conducted three to five weeks before each primary) and 4.3 percentage points (for respondents deemed "most likely" to vote) in final polls (conducted the weekend before each primary). We also found that Gallup consistently underestimated support for Pat Robertson and Jesse Jackson, in each case by about five percentage points.

[24] We have seen that with media tracking polls, some of the confusion stems from the inevitable but often unacknowledged time lag between polling and publication. In general, tracking polls seem more useful for campaign planning (and in some instances, perhaps, for historical analysis by academics) than for campaign coverage by electronic or, especially, print media.

[25] This is one explanation for the fact that so much of the analysis of the CBS News/*New York Times* polls involved cross-tabulations rather than marginal frequencies. Recall from Table 3 that the percentage of poll stories involving cross-tabular analysis of one kind or another was more than 90% for CBS News/*New York Times* surveys, about 60% for Gallup, and only 2% for Harris.

[26] The impact of primary outcomes on subsequent campaign events is documented in Bartels, 1988. Analyses of poll effects more specifically include, for example, Beniger, 1976; Sudman, 1986; and Skalaban, 1988.

The pollsters' treatment of Jesse Jackson illustrates the importance of interpreting rather than simply reporting raw levels of candidate support. Unquestionably, Jackson evoked negative sentiment among a segment of the population. Some of this negative sentiment resulted from his political experience, policy positions, campaign statements, and acquaintances; but some of it was also due to "old fashioned racism" (McConahay, 1986).[27] Pollsters documented the negative image of Jackson, but they provided precious little of the information necessary to analyze or account for that image.[28]

Conversely, pollsters helped to document and publicize the popularity of Michael Dukakis at a point in the campaign season when Dukakis was far from assured of the Democratic nomination. As with Jackson, we find ourselves wondering whether the polls contributed to Dukakis's emergence by telling people *why* he was popular, or simply by telling people *that* he was popular. The first enterprise is harder, and therefore rarer—but in our view more valuable.

Much of the tension inherent in the pollsters' role, in this as in other electoral settings, stems from the fact that their public is simultaneously a passive audience and an active citizenry. To the extent that pollsters focus on the "horse race"—who is ahead, who won the debate, who is losing ground—they may provide an exciting story while estranging the public from the real politics of the electoral process. Elections are supposed to be plebiscites on the parties, the issues, the candidates, and the management of government. The pollsters' fixation on the horse race helps to turn presidential selection into something more like a sporting event—a yearlong Kentucky Derby with citizens as the spectators.

Media coverage of the presidential nominating process sometimes suggests that the political significance of a primary or caucus is to be measured by its influence on the candidates' standings in the national opinion polls. The early events in Iowa and New Hampshire are considered important because the winners in those contests tend to rise in the polls and the losers tend to fall. It is as if voters in these and other states went to the polls to record their preferences in order to produce a "bounce" among 1,500 anonymous strangers—the respondents in the next week's poll. It would be well for pollsters and poll analysts alike to bear in mind that, for better or worse, poll results (and interpretations of those results) are not the end of the story. Polls are important because they, in turn, influence prospective voters. Presidents are still nominated and elected not *by* the polls, but *at* the polls.

---

[27] In a 1984 Gallup Poll, 23% of the population said they would not vote for a black candidate for president even if the candidate were "qualified." Presumably, these people's opinions would have been "unfavorable" or "very unfavorable" toward Jesse Jackson in 1988 when he emerged as a contender for the nomination. See Walters, 1988.

[28] Some Harris questions of the sort shown in Table 9 were the most notable exception to this pattern.

References

Bartels, Larry M. (1988)
Presidential Primaries and the Dynamics of Public Choice. Princeton: Princeton University Press.

Beniger, James R. (1976)
"Winning the presidential nomination: National polls and state primary elections, 1936–1972." Public Opinion Quarterly 40:22–38.

Broder, David S. (1988)
"The Jackson message." Washington Post, 8 March.

Broh, C. Anthony (1980)
"Horse-race journalism: Reporting the polls in the 1976 presidential election." Public Opinion Quarterly 44:514–529.

——— (1987)
A Horse of a Different Color: Television's Coverage of Jesse Jackson's 1984 Presidential Campaign. Washington: Joint Center for Political Studies.

Crespi, Irving (1988)
Pre-Election Polling: Sources of Accuracy and Error. New York: Russell Sage Foundation.

Dionne, E. J., Jr. (1988)
"Poll shows Dukakis leads Bush: Many Reagan backers shift sides." New York Times, 17 May.

Greenfield, Meg (1988)
"A new kind of racial putdown." Newsweek, 4 April, 80.

Hamilton, Arnold (1988)
"Jackson on ticket damages Dukakis, new survey shows." San Jose Mercury News, 28 April.

Harris, Louis (1988)
"The strange wild cards in the 1988 election: Jackson, Hart, Robertson." Press release, 7 January.

Massey, James T., and Steven L. Botman (1988)
"Weighted adjustments for random digit dialed surveys." Pp. 143–160 in Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, and Joseph Waksberg, Telephone Survey Methodology. New York: Wiley.

McConahay, John B. (1986)
"Modern racism, ambivalence, and the modern racism scale." Pp. 91–125 in John Dovidio and Samuel Gaertner (eds.), Prejudice, Discrimination, and Racism. Orlando: Academic Press.

Mitofsky, Warren J. (1981)
"The 1980 pre-election polls: A review of disparate methods and results." Paper presented at the annual meeting of the American Statistical Association, Detroit, 10–13 August.

Rosenthal, Andrew (1988)
"Accuracy of surveys fails to sway some critics." New York Times, 10 November.

Shapiro, Walter (1988)
"Win, Jesse, win: After Michigan, the Democrats have an unexpected front runner." Time, 4 April, 21.

Skalaban, Andrew (1988)
"Do the polls affect elections? Some 1980 evidence." Political Behavior 10: 136–150.

Sudman, Seymour (1986)
"Do exit polls influence voting behavior?" Public Opinion Quarterly 50:331–339.

Taylor, Paul (1988)
"Is media polling messing up the primaries?" Washington Post, national weekly edition, 7–13 March.
Thornberry, Owen T., and James T. Massey (1988)
"Trends in United States telephone coverage across time and subgroups." Pp. 25–50 in Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, and Joseph Waksberg, Telephone Survey Methodology. New York: Wiley.
Walters, Ronald W. (1988)
Black Presidential Politics in America. Albany: State University of New York Press.

THE POLLS—A REPORT
1980–1988 NEW HAMPSHIRE PRESIDENTIAL PRIMARY POLLS

KENNETH E. JOHN

The New Hampshire primary remains a pivotal event in the presidential campaign. Vice President George Bush overcame a dismal third-place finish in the 1988 Iowa caucus to surge to victory in New Hampshire, dealing a fatal blow to the campaign of Sen. Robert Dole. Gov. Michael Dukakis's victory in the 1988 New Hampshire primary, following his third-place finish in Iowa, cemented his position as the Democratic frontrunner.

In 1984 Sen. Gary Hart was soundly beaten in the Iowa caucus, but his surprise finish in second place catapulted him into the limelight, and his upset win in New Hampshire shortly thereafter led to a sustained challenge for the nomination. In 1980 Ronald Reagan overcame defeat in Iowa to beat George Bush in the Granite State.

Tracking polls and other surveys have chronicled major developments in every seriously contested New Hampshire presidential primary in the 1980s. Although private tracking surveys were conducted in 1980, the first public tracking poll gauging day-by-day changes in New Hampshire preference was conducted by the *Washington Post* and ABC News in 1984.

For the purposes of this article, "tracking polls" are defined as polls in which fieldwork is conducted every day in the week prior to election day and in which results are analyzed and/or reported several times over that period. Reporting of such "tracking polls" can be based on figures covering a single day, more than one day, or on overlapping rolling averages with each covering multiple days of interviewing. A compilation of New Hampshire presidential primary polls follows this article, including data from tracking polls and other surveys conducted in the 1980s. Final election results are also presented.

The main objective of tracking polling before the New Hampshire

KENNETH E. JOHN is a Senior Analyst at Schulman, Ronca, & Bucuvalas, Inc.

Public Opinion Quarterly Volume 53:590–605 © 1989 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/89/0053-04/$2.50

primary is to monitor changes in electoral preference. Tracking polls continue as late as possible in the campaign, sometimes including interviews on the evening before primary day. By polling as close to the primary as possible, tracking polls essentially attempt to predict the actual vote. Previous experience has shown that preprimary polls can easily be dated, particularly those completed more than a week before the primary is held (Dionne, 1980; Felson and Sudman, 1975). Tracking polls overcome this limitation. One recent analysis has found that the time between a preelection poll and the election is one of the most important variables associated with the accuracy of the survey (Crespi, 1988).

As with any other preprimary polls, the task of tracking polls is difficult, in part because primary turnout in New Hampshire and other states comprises only a minority of eligible registered voters. Little research has been published on questions used in preprimary likely-voter models along the lines of Paul Perry's (1979, 1973) analysis of national surveys or subsequent analysis of simplified likely-voter models (Traugott and Tucker, 1984). Thus, construction of "likely voter" weights for preprimary tracking polls appears to be a matter of individual judgment by each polling organization.

New Hampshire tracking polls to date share other drawbacks that apply to tracking polls in general. First, the length of the interview is truncated to increase the number of completed interviews, limiting the depth of questioning. Indeed, tracking polls almost always emphasize "horserace" reporting, perhaps to the detriment of reporting on more stable, basic dimensions of public opinion such as campaign issues or voter motivation. Second, tracking poll sample sizes are relatively small, particularly if the survey is limited to a subset of adults deemed to be likely voters. Analysis of shifting preferences of subgroups is therefore limited because several days' interviewing must be combined to obtain sufficient sample sizes for the subgroup of interest. Third, while tracking polls can be based on a mixture of first attempts and callbacks, the difficulty of finding random samples on particular days of the week, such as Fridays and Saturdays, may limit the validity of tracking poll findings for those days. Fourth, time for analysis of tracking polls is limited because of the need for haste in reporting the figures in a timely manner. Fifth, the predictive accuracy of tracking polls has been mixed.

It is possible that the tracking polls reveal change in opinion, often suggesting volatility of preference. An alternative explanation, however, is that many voters are poorly informed or have not focused on the contest, so that the tracking poll numbers shift as people make up their minds.

Despite these disadvantages, there are several reasons many public opinion polling organizations find tracking polling prior to the New Hampshire primary to be worthwhile, if not compelling. Tracking polls provide an independent source of data on the "horserace," removing the need to rely on campaign polls and anectodal sources of information.

Consider the 1980 New Hampshire Republican primary. News media reports credited Reagan's 1980 New Hampshire primary triumph to his success in a Nashua, New Hampshire, debate with George Bush. However, Richard Wirthlin, who conducted private tracking polls for Reagan, said an earlier debate in Manchester was the turning point of the campaign (Germond and Witcover, 1981), a point of view later shared by some Bush aides (Harwood, 1980) and journalists (Broder, 1987). If this interpretation is correct, the media initially missed the significance of the Manchester debate and seriously distorted the true significance of the crucial Nashua debate. Such considerations encouraged news organizations to begin their own tracking polls.

Aside from giving reporters and the public access to current findings, public tracking polls have other advantages. They provide frequent, up-to-date information on voter attitudes and preference that complements other reporting efforts and helps shape news coverage. While sample sizes are limited, some organizations are able to conduct sufficient interviews to reduce sampling error to tolerable levels for the subset of respondents classified as likely voters. The use of overlapping rolling averages further increases the available sample sizes for reporting purposes. In addition, major polling organizations have found that the time constraints are not so stringent as to preclude analysis of the main findings.

By polling every day in a specified period before the primary, tracking polls aid in the identification of key turning points in the primary campaign. Opinion can be analyzed before and after key intervening events, such as debates. Tracking polls thus provide a unique source of information on changes in voter preferences, a vital advantage in New Hampshire, where tidal shifts in preference have not been uncommon in the final days before primaries.

For example, the *Washington Post* and ABC News tracking polling in the New Hampshire Democratic primary in 1984 was the first time a major news media organization had conducted a tracking poll in the Granite State. The poll found Hart moving into second place after the Iowa caucus and increasing his support from 23% to 35% over the final three days of the campaign. Mondale dropped from 36% to 27% over the same period (Begans, 1985). Hart received 37% of the actual votes cast, and Mondale received 28%. The *Post* had only reported the average of the final three days of interviewing, showing Hart moving into a tie. Still, it was the only public poll to suggest Hart would be victorious.

Later both campaigns revealed findings from their private tracking polls, in which interviews were conducted up to the Saturday before the primary. Both showed a surge in positive attitudes toward Hart, which campaign pollsters interpreted as a precursor to a shift in preference that did in fact occur in the final days of the campaign (Moore, 1986).

Many organizations decided to conduct New Hampshire polls in 1988. Three polling organizations found similar patterns of voter preference in the hotly contested GOP primary, but they had widely varying findings in their final figures (see Fig. 1).

The final CBS News tracking poll before the February 16 GOP primary found that "Bush's slim lead (4 points on February 14–15) is within the margin of error and comes entirely from those Republicans and independents who say they will 'probably vote' in the Republican primary" (CBS News Poll, 1988). The final *Post*/ABC News tracking poll on February 15 showed a tie. The final Gallup data, based on interviews conducted February 12–14, showed Dole with a 5-point margin among likely voters and an 8-point lead among those most likely to vote. The final *Boston Globe* tracking poll, taken February 13–14 and published February 15, listed Bush with 31% and Dole with 30%. The *Globe*, however, ran a correction on February 28 stating that the final poll figures should have been Bush 38%, Dole 29%. A February 14 poll by the *Boston Herald* and WBZ-TV found Dole with 32% and Bush with 31%. Bush won the primary with 38%, while Dole placed second with 28%.

In the Democratic race, Dukakis led by a wide margin during 1988 and won the primary with 36%. Much media attention focused on which candidate would finish second. Representative Richard Gephardt, who won the Iowa caucus, eventually edged Senator Paul Simon by 20% to 17%. Again, the record of the polls was mixed. The *Herald*/WBZ-TV poll showed a 6-point Gephardt edge over Simon; Gallup showed a 4-point Gephart margin; the *Washington Post*/ABC News tracking poll showed a tie; the *Globe* tracking poll showed a 3-point Simon edge; and the CBS News tracking poll also showed a 3-point Simon margin.

After the primary, Gallup distributed a memo to clients that said they had not placed enough emphasis on how preferences could change overnight and on the softness in Dole's support. The *Post*/ABC News tracking poll also did not give enough emphasis to differences in the strength of support (Bush's supporters were more likely than Dole's backers to call themselves "strong" supporters). Later *Post*/ABC News tracking polls gave greater weight to "strong supporters" in

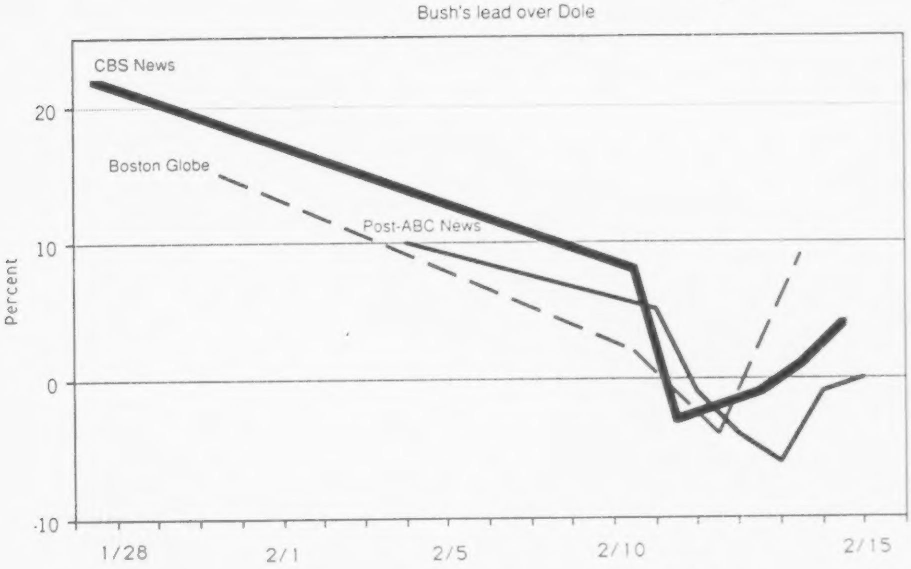


Figure 1. New Hampshire Polls: The 1988 GOP Primary

tabulating the horserace and used a tighter screening process for likely voters. These examples illustrate the difficulty of conducting accurate New Hampshire primary polls and show the evolving nature of tracking polls in that state.

Despite their failings, the tracking polls did make a real contribution to understanding the 1988 GOP campaign. All the surveys showed that Dole and Gephardt received sizable boosts in support immediately after the Iowa Republican and Democratic caucuses. The published polling data by CBS News, the *Post*/ABC News, and the *Boston Globe* showed Dole moving slightly ahead, but with Bush staging a dramatic reversal of the trend in the closing days of the campaign. Figure 1 shows published data from these three surveys.

The CBS News New Hampshire exit poll was consistent with the tracking polls. Bush had the lead among voters who made up their minds before the Iowa caucus, while Dole won a plurality among voters who decided on a candidate shortly after the Iowa caucus. The late deciders, i.e., those who made up their minds in the three days preceding the primary, voted for Bush over Dole by 36% to 30%.

For better or worse, media coverage of presidential primaries has changed, perhaps forever, as a result of immediate feedback from tracking surveys. Expectations are geared to the most up-to-date tracking data rather than to older results or no results at all. For a few brief days in 1988, Dole was regarded as the new frontrunner in the GOP campaign. His defeat was regarded as a great disappointment relative to his high-water mark reported by tracking polls.

In devising tactics, the campaigns appear to be relying on public and private tracking surveys more than ever. When Dole took the lead in New Hampshire tracking surveys, he appeared to become overconfident. "I felt pretty good for about 24 hours," Dole said. "I was writing my inaugural address" (Schribman, 1988). Dole did not take steps to widely air advertisements with a no-tax-increase pledge, an omission that may have been a costly mistake. In contrast, after public tracking polls showed the vice president losing support, the Bush campaign made a series of changes, including "attack ads" and shifting to a more personal style of campaigning.

The tracking polls, according to one theory, prompted a countertrend for Bush when media accounts portrayed Dole as headed for victory, encouraging strategic voting to help keep Bush in the race (Taylor, 1988). Little empirical evidence appears available to test this theory, which was advocated perhaps most vigorously by Dole campaign officials. In general, there is little polling data that can be used to measure the effect of media reporting of tracking polls in New Hampshire.

However, tracking surveys have helped shape interpretation of New Hampshire primary politics. By showing that preferences shifted sharply before the 1984 and 1988 New Hampshire primaries, tracking polls gave support to David Moore's (1984) contention that "retail politics" involving one-on-one contact between the candidate and the voter is not the determining factor and can be overwhelmed by the effects of media reports following the Iowa caucuses.

News media organizations are increasing their reliance on tracking polls to gauge New Hampshire primary preference. The tracking polls, for all their failings, are attractive because they can be conducted frequently and provide an up-to-date barometer of sentiment. They have proved to be indispensible, if not infallible, to the news media interpretation of recent New Hampshire presidential primaries. The lessons of the tracking polling to date: poll up to the final day before the election, screen for likely voters, and take into account the strength of support.

The survey information reported here was drawn from published and unpublished material with the assistance of the Roper Center and most of the survey organizations. Further information about this data can be obtained from the appropriate organization or archive. Surveys classified as tracking polls are marked with a *T*. Figures are for "likely voters" unless otherwise indicated.

CBS—Columbia Broadcasting System
CBS/NYT—Columbia Broadcasting System/*New York Times*
CNN/DRC—Cable News Network, conducted by Darden Research Corporation

Gallup—The Gallup Organization
Globe—*Boston Globe*
Herald/WBZ—*Boston Herald*/WBZ Television
LAT—*Los Angeles Times*
NBC—National Broadcasting Company
WCVB/WHDH—WCVB Television, Needham, and WHDH Radio, Boston
WP/ABC—*Washington Post*/American Broadcasting Company

References

Begans, Peter (1985)
"A roller coaster ride through the 1984 Democratic primaries." Pp. 571–577 in The '84 Vote. New York: ABC News.
Broder, David S. (1987)
Behind the Front Page. New York: Simon and Schuster.
CBS News Poll (1988)
"CBS News New Hampshire primary poll—5." CBS News Poll news release, 16 February.
Congressional Quarterly (1988)
"Opening round primaries and caucuses . . . fragmented electorate in many contests." 5 March, 570–571.
Crespi, Irving (1988)
Pre-Election Polling: Sources of Accuracy and Error. New York: Russell Sage Foundation.
Dionne, E. J. (1980)
"New Hampshire poll backs view of volatile mood among voters." New York Times, 2 March, 1.
Felson, Marcus, and Seymour Sudman (1975)
"The accuracy of presidential-preference primary polls." Public Opinion Quarterly 39:232–236.
Germond, Jack W., and Jules Witcover (1981)
Blue Smoke and Mirrors. New York: Viking Press.
Harwood, Richard, ed. (1980)
The Pursuit of the Presidency 1980. New York: Berkeley Books.
Moore, David (1984)
"The death of politics in New Hampshire." Public Opinion 7(1):56–57.
Moore, Jonathan, ed. (1986)
Campaign for President. Dover: Auburn House.
Perry, Paul (1973)
"A comparison of the voting preferences of likely voters and likely nonvoters." Public Opinion Quarterly 37:99–109.
——— (1979)
"Certain problems in election survey methodology." Public Opinion Quarterly 43:312–325.
Schribman, David (1988)
"Bungled campaign: How Dole squandered money, opportunities, dooming hopes for '88." Wall Street Journal, 14 March, 1.
Scammon, Richard M., and Alice V. McGillivray (1981)
America Votes 14. Washington: Congressional Quarterly Elections Research Center.

——— (1985)
American Votes 16. Washington: Congressional Quarterly Elections Research Center.
Taylor, Paul (1988)
"Did pollsters give Bush his N.H. win?" Washington Post, 28 February, C2.
Traugott, Michael W., and Clyde Tucker (1984)
"Strategies for predicting whether a citizen will vote and estimation of electoral outcomes." Public Opinion Quarterly 48:330–343.

1988 Democratic Primary

CBS (T): Who do you want the Democrats to nominate for President in 1988—Bruce Babbitt, Michael Dukakis, Dick Gephardt, Albert Gore, Gary Hart, Jesse Jackson or Paul Simon? (IF "UNDECIDED" OR "NO ONE" ASK:) As of today, do you lean more towards Bruce Babbitt, Michael Dukakis, Dick Gephardt, Albert Gore, Gary Hart, Jesse Jackson or Paul Simon?

	1/25–30	*2/9–10*	*2/10–11*	*2/12–13*	*2/13–14*	*2/14–15*
Babbitt	3%	2%	2%	3%	3%	4%
Dukakis	40	38	39	42	37	38
Gephardt	7	19	16	18	15	13
Gore	4	7	7	4	4	3
Hart	13	4	5	4	4	4
Jackson	6	4	5	4	8	8
Simon	16	15	14	12	14	16
Undecided	11	11	12	13	15	14
	100	100	100	100	100	100
N	434	493	466	483	475	479

Gallup: I will read you the names of the seven Democratic Presidential candidates. After I read all seven names, please tell me which one you would like to see nominated as the Democratic Party's candidate for president. Bruce Babbitt, Michael Dukakis, Richard Gephardt, Albert Gore, Gary Hart, Jesse Jackson, Paul Simon. (*Note:* names were rotated.) (IF "UNDECIDED," ASK:) As of today, to which of the Democratic candidates do you lean most?

	1/8–10	*2/12–13*	*2/12–14*	*2/12–13*[a]	*2/12–14*[a]
Babbitt	3%	3%	4%	5%	5%
Dukakis	39	42	43	39	39
Gephardt	5	18	17	18	18
Gore	4	4	5	5	6
Hart	19	4	4	3	3
Jackson	5	6	6	7	7
Simon	12	14	12	16	14
Undecided/other	13	9	9	7	8
	100	100	100	100	100
N	416	756	1,177	424	644

[a] Most likely voters.

Globe (T) [Question wording not available]

	1/27–2/3	*2/9–10*	*2/11–12*	*2/13–14*
Babbitt	6%	3%	5%	3%
Dukakis	38	35	38	35
Gephardt	9	18	16	12
Gore	6	5	3	5
Hart	11	3	3	5
Jackson	7	8	9	6
Simon	17	12	14	15
Undecided	6	15	11	18
	100	100	100	100
N	600	329	NA	NA

Herald/WBZ: The candidates for president in the Democratic presidential primary in New Hampshire are Bruce Babbitt, Michael Dukakis, Richard Gephardt, Albert Gore, Gary Hart, Jesse Jackson and Paul Simon. If the election were held today, for whom would you vote? (IF "UNDECIDED," ASK:) Is there any candidate you are leaning toward voting for? (REPEAT LIST IF NECESSARY)

	1/25–26	*2/9*	*2/14*
Babbitt	5%	4%	3%
Dukakis	35	44	38
Gephardt	7	17	19
Gore	4	3	4
Hart	10	5	4
Jackson	5	6	6
Simon	9	13	13
Undecided	25	9	12
	100	100	100
N	402	400	400

LAT: If the February 16th primary for President were being held today, who, if anyone, would you vote for, or haven't you heard enough about that yet to say? (IF "HAVEN'T HEARD ENOUGH," "UNDECIDED" OR "REFUSED," ASK:) Well, suppose these were the Democratic candidates for President: former Arizona Governor Bruce Babbitt, or Massachusetts Governor Michael Dukakis, or Missouri Congressman Richard Gephardt, or Tennessee Senator Albert Gore, or the Reverend Jesse Jackson, or Illinois Senator Paul Simon, which, if any, of these candidates would you vote for, or haven't you heard enough about that yet to say?

	1/19–22
Babbitt	13%
Dukakis	37
Gephardt	8
Gore	5
Hart	6

Jackson	6
Simon	18
Undecided	6
	100
N	487

WP/ABC (T): Imagine that the New Hampshire Democratic primary were being held today and these are the candidates: Bruce Babbitt, Michael Dukakis, Richard Gephardt, Albert Gore, Gary Hart, Jesse Jackson and Paul Simon. For whom would you vote: Babbitt, Dukakis, Gephardt, Gore, Hart, Jackson or Simon? (*Note:* names were rotated in the last sentence.) (IF "UNDECIDED" OR "REFUSED," ASK:) Which one would you say you are leaning to right now?

	2/1–7	*2/9*	*2/9–11*	*2/10–12*	*2/11–12*	*2/12–14*	*2/13–15*
Babbitt	6%	4%	4%	4%	6%	6%	5%
Dukakis	43	37	37	40	40	41	41
Gephardt	12	21	19	16	16	16	17
Gore	5	5	6	7	7	7	7
Hart	7	4	3	3	3	5	5
Jackson	6	5	7	8	6	6	5
Simon	13	17	17	15	16	16	17
Undecided/ other	8	7	7	7	6	3	3
	100	100	100	100	100	100	100
N	503	403	707	456	455	708	963

	2/14	*2/15*
Babbitt	6%	3%
Dukakis	41	47
Gephardt	15	16
Gore	7	6
Hart	5	5
Jackson	8	7
Simon	16	15
Undecided/other	2	1
	100	100
N	403	408

NOTE: *Washington Post* figures. Slight differences in the way the *Post* and ABC News weighted the sample resulted in minor differences in results reported by each organization.

1988 Republican Primary

CBS (T): Who do you want the Republicans to nominate for president in 1988—George Bush, Bob Dole, Pete du Pont, Alexander Haig, Jack Kemp or Pat Robertson? (IF "UNDECIDED" OR "NO ONE" ASK:) As of today, do you lean more towards George Bush, Bob Dole, Pete du Pont, Alexander Haig, Jack Kemp or Pat Robertson?

	1/25–30	*2/9–10*	*2/10–11*	*2/12–13*	*2/13–14*	*2/14–15*
Bush	42%	35%	29%	30%	32%	34%
Dole	20	27	32	31	31	30
du Pont	6	7	7	10	9	10
Haig	1	2	1	[a]	[a]	[a]
Kemp	16	12	12	11	9	9
Robertson	4	10	10	10	9	8
Undecided	11	7	9	8	10	9
	100	100	100	100	100	100
N	513	543	499	594	588	596

[a] Not included on the list of candidates.

Gallup: I will read you the names of the six Republican Presidential candidates. After I read all six names, please tell me which one you would like to see nominated as the Republican Party's candidate for president. George Bush, Robert Dole, Pete du Pont, Alexander Haig, Jack Kemp, Pat Robertson. (*Note:* names were rotated.) (IF "UNDECIDED," ASK:) As of today, to which of the Republican candidates do you lean most?

	1/8–10	*2/12–13*	*2/12–14*	*2/12–13*[a]	*2/12–14*[a]
Bush	38%	28%	28%	28%	27%
Dole	23	33	33	36	35
du Pont	5	9	9	7	8
Haig	4	[b]	[b]	[b]	[b]
Kemp	15	12	12	12	13
Robertson	4	11	10	11	10
Undecided/other	11	7	8	6	7
	100	100	100	100	100
N	545	993	1,531	589	902

[a] Most likely voters.

[b] Not included on the list of candidates.

Globe (T) [Question wording not available]

	1/27–2/3	*2/9–10*	*2/11–12*	*2/13–14*
Bush	39%	30%	28%	38%
Dole	19	28	32	29
du Pont	8	8	10	[a]
Haig	3	1	2	[a]
Kemp	17	12	8	[a]
Robertson	7	11	11	[a]
Undecided	7	10	9	[a]
	100	100	100	
N	600	500	[a]	298

[a] Information not available.

Herald/WBZ: The candidates for president in the Republican presidential primary in New Hampshire are George Bush, Robert Dole, Pierre du Pont, Alexander Haig, Jack Kemp and Pat Robertson. If the election were held today, for whom would you vote? (IF "UNDECIDED" ASK:) Is there any candidate you are leaning toward voting for? (REPEAT LIST IF NECESSARY)

	1/27	*2/9–10*	*2/14*
Bush	40%	32%	31%
Dole	18	23	32
du Pont	5	9	8
Haig	2	2	0
Kemp	11	11	9
Robertson	3	9	8
Undecided	21	14	11
	100	100	100
N	[a]	400	525

[a] Information not available.

LAT: If the February 16th primary for President were being held today, who, if anyone, would you vote for, or haven't you heard enough about that yet to say? (IF "HAVEN'T HEARD ENOUGH," "UNDECIDED" OR "REFUSED," ASK:) Well, suppose these were the Republican candidates for president: Vice President George Bush, and Kansas Senator Robert Dole, and former Delaware Governor Pete du Pont, and former Secretary of State Alexander Haig, and New York Congressman Jack Kemp, and T.V. Evangelist Pat Robertson, which, if any, would you vote for, or haven't you heard enough about that yet to say?

	1/19–22
Bush	35%
Dole	22
du Pont	8
Haig	2
Kemp	18
Robertson	8
Undecided	7
	100
N	675

NBC: If the New Hampshire Republican presidential primary were being held today, would you vote for: George Bush, Robert Dole, Pete du Pont, Alexander Haig, Jack Kemp or Pat Robertson? (*Note:* names were rotated)

	1/17–19
Bush	37%
Dole	22
du Pont	7
Haig	2
Kemp	11

Robertson	5
Someone else (vol.)	1
Undecided	15
	100
N	946

NOTE: Reinterviews from an NBC poll of 1,145 Republican likely voters conducted 15–21 November 1987.

WP/ABC (T): Imagine that the New Hampshire Republican primary were being held today and these are the candidates: George Bush, Bob Dole, Pierre du Pont, Al Haig, Jack Kemp and Pat Robertson. For whom would you vote: Bush, Dole, Du Pont, Haig, Kemp and Robertson? (*Note:* names were rotated in the last sentence) (IF "UNDECIDED" OR "REFUSED," ASK:) Which one would you say you are leaning to right now?

	2/1–7	*2/9*	*2/9–11*	*2/10–12*	*2/11–13*	*2/12–14*	*2/13–15*
Bush	36%	33%	32%	29%	28%	28%	30%
Dole	26	29	27	30	32	34	33
du Pont	9	7	10	11	11	10	10
Haig	3	2	2	1	1	1	0
Kemp	13	12	12	12	13	14	13
Robertson	6	9	12	11	10	9	8
Undecided/ other	7	8	5	6	5	4	6
	100	100	100	100	100	100	100
N	501	404	712	467	457	718	969

	2/14	*2/15*
Bush	32%	32%
Dole	33	32
du Pont	8	10
Haig	0	1
Kemp	14	12
Robertson	8	8
Undecided/Other	5	5
	100	100
N	402	405

NOTE: *Washington Post* figures. Slight differences in the way the *Post* and ABC News weighted the sample resulted in minor differences in results reported by each organization.

1984 Democratic Primary

WP/ABC (T): Imagine that the New Hampshire Democratic primary is being held today and these are the candidates: Reubin Askew, Alan Cranston, John

Glenn, Gary Hart, Ernest Hollings, Jesse Jackson, George McGovern and Walter Mondale. For whom would you vote: Askew, Cranston, Glenn, Hart, Hollings, Jackson, McGovern or Mondale?

	2/13–20	*2/21–23*	*2/24–26*	*2/25–27*
Hart	13%	22%	24%	30%
Glenn	20	14	14	14
Jackson	10	7	6	6
Mondale	37	39	38	30
Undecided/other	20	18	18	20
	100	100	100	100
N	526	343	351	446

	2/25	*2/26*	*2/27*
Hart	23%	33%	35%
Mondale	36	26	27
Undecided/other	41	41	38
	100	100	100
N	150[a]	150[a]	150[a]

[a] Approximate sample size.

CNN/DRC [Question wording not available; makeup of sample not known]

	2/25–26
Hart	32%
Glenn	15
Jackson	7
Mondale	38
Undecided/other	8
	100
N	500

Globe [Question wording and sample sizes not available]

	1/29–31	*2/12–14*
Askew	1%	1%
Cranston	2	2
Hart	12	13
Hollings	2	2
Glenn	18	16
Mondale	37	36
Jackson	16	10
McGovern	4	6
Undecided/other	8	14
	100	100

WCVB/WHDH [Question wording and sample size not available]

	2/21–22
Hart	13%
Mondale	39
Glenn	13
Jackson	7
Undecided/other	28
	100

1980 Democratic Primary

CBS/NYT: Who would you like to see nominated as the Democratic candidate for president in 1989—Jerry Brown, Jimmy Carter, or Edward Kennedy?

	2/22–24[a]	*2/27–28*[b]
Brown	4%	6%
Carter	50	55
Kennedy	30	26
No opinion/other	16	13
	100	100
N	470	414

[a] Includes registered Democrats and independents who said they would vote in the Democratic primary.

[b] Postelection survey based on registered voters reinterviewed from 2/22–24 survey.

1980 Republican Primary

CBS/NYT: Who would you like to see nominated as the Republican candidate for president in 1980—John Anderson, Howard Baker, George Bush, John Connally, Philip Crane, Robert Dole, Ronald Reagan, or someone else?

	2/22–24[a]	*2/27–28*[b]
Anderson	7%	6%
Baker	13	11
Bush	28	13
Connally	3	1
Crane	1	1
Dole	1	0
Reagan	38	58
No opinion/other	9	10
	100	100
N	597	505

[a] Includes registered Republicans and independents who said they would vote in the Republican primary.

[b] Postelection survey based on registered voters reinterviewed from 2/22–24 survey.

Globe [Question wording not available]

	1/24–27[a]	*1/24–27*
Bush	45%	43%
Reagan	36	39
Undecided/other	19	18
	100	100
N	616	275

[a] Registered voters.

New Hampshire Presidential Primary: Final Results

1988 (16 FEBRUARY)

Democratic		*Republican*	
Michael S. Dukakis	35.8%	George Bush	37.6%
Richard A. Gephardt	19.8	Robert J. Dole	28.4
Paul Simon	17.1	Jack Kemp	12.8
Jesse L. Jackson	7.8	Pierre S. (Pete) du Pont IV	10.1
Albert Gore, Jr.	6.8	Marion G. (Pat) Robertson	9.4
Bruce Babbitt	4.6	Alexander M. Haig, Jr.	0.3
Gary Hart	4.0	Others	1.4
Others	4.1		100.0
	100.0		

1984 (28 FEBRUARY)

Democratic		*Republican*	
Gary Hart	37.3%	Ronald Reagan	86.1%
Walter F. Mondale	27.9	Others	13.9
John Glenn	12.0		100.0
Jesse L. Jackson	5.3		
George McGovern	5.2		
Ernest F. Hollings	3.5		
Alan Cranston	2.1		
Reubin Askew	1.0		
Others	5.8		
	100.0		

1980 (26 FEBRUARY)

Democratic		*Republican*	
Jimmy Carter	47.1%	Ronald Reagan	49.6%
Edward M. Kennedy	37.3	George Bush	22.7
Edmund G. (Jerry) Brown	9.6	Howard H. Baker, Jr.	12.9
Others	6.0	John B. Anderson	9.8
	100.0	Philip M. Crane	1.8
		John B. Connally	1.5
		Robert J. Dole	0.4
		Others	1.3
			100.0

10610-1

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

R. DARCY and MICHAEL LAVER	Referendum Dynamics and the Irish Divorce Amendment
TOM W. SMITH	The Origins of Election Polls
LEE SIGELMAN	Reagan, the Public, and the INF Treaty
STEPHEN ERFLE HENRY MCMILLAN and BERNARD GROFMAN	Regulation via Threats: Politics, Media Coverage, and Oil Pricing Decisions
BRUCE C. STRAITS	The Social Context of Voter Turnout
JACK CITRIN DONALD PHILIP GREEN and DAVID O. SEARS	White Reactions to Black Candidates
ERIC R. A. N. SMITH and PEVERILL SQUIRE	Prestige Names in Question Wording
THE POLLS	Review: The Strange Career of the "Harvard-Yale Study" *by Andrew Cherlin*
	Report: Privacy and Information Technology *by James E. Katz and Annette R. Tassone*
BOOKS	Reviews and Notes *by Michael MacKuen, Roy Pierce, Kurt Lang, Michael W. Traugott, and Shoshana Colman*

Spring 1990

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) David L. Sills (1986–90)
Gladys Engel Lang (1989–93)

Norman Bradburn and Andrew Kohut,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor David Carlson
Office Manager Carlotta Nosel
Assistant to the Editor Cheryl Rieger

The Polls
Reviews: Stanley Presser, Seymour Sudman
Reports: Tom W. Smith

EDITORIAL BOARD

Robert P. Abelson
Barbara A. Bailar
Sandra J. Ball-Rokeach
Larry M. Bartels
George F. Bishop
Lawrence D. Bobo
Jean M. Converse
Thomas D. Cook
Irving Crespi
James A. Davis
Stanley Feldman
Richard J. Fox
Norval D. Glenn
Robert M. Groves
Deborah R. Hensler
Shanto Iyengar
Jon A. Krosnick
Kurt Lang
Michael MacKuen
McKee J. McClendon
Jack M. McLeod
John Modell
John Mueller
Harold Nisselson
Helmut Norpoth
Steven J. Rosenstone
Norbert Schwarz
David O. Sears
Robert Y. Shapiro
Lee Sigelman

Public Opinion Quarterly (ISSN 0033-362X) Volume #54, 4 issues (1990)

© 1990 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 54 / SPRING 1990 / NUMBER 1

Articles

Referendum Dynamics and the Irish Divorce Amendment 1
R. DARCY and MICHAEL LAVER

The First Straw? A Study of the Origins of Election Polls 21
TOM W. SMITH

Disarming the Opposition: The President, the Public, and the INF Treaty 37
LEE SIGELMAN

Regulation via Threats: Politics, Media Coverage, and Oil Pricing Decisions 48
STEPHEN ERFLE, HENRY MCMILLAN, and BERNARD GROFMAN

The Social Context of Voter Turnout 64
BRUCE C. STRAITS

White Reactions to Black Candidates: When Does Race Matter? 74
JACK CITRIN, DONALD PHILIP GREEN, and DAVID O. SEARS

The Effects of Prestige Names in Question Wording 97
ERIC R. A. N. SMITH and PEVERILL SQUIRE

The Polls

Review: The Strange Career of the "Harvard-Yale Study" 117
ANDREW CHERLIN

Report: Public Opinion Trends—Privacy and Information Technology 125
JAMES E. KATZ and ANNETTE R. TASSONE

Books

David P. Fan, *Predictions of Public-Opinion from the Mass Media: Computer Content Analysis and Mathematical Modeling* 144
MICHAEL MACKUEN

Donald Granberg and Sören Holmberg, *The Political System Matters: Social Psychology and Voting Behavior in Sweden and the United States* 147
ROY PIERCE

Richard V. Ericson, Patricia M. Baranek, and Janet B. L. Chan, *Negotiating Control: A Study of News Sources* 149
KURT LANG

Norman M. Bradburn and Seymour Sudman, *Polls and Surveys: Understanding What They Tell Us* 152
MICHAEL W. TRAUGOTT

Book Note: Thomas F. Pettigrew and Denise A. Alston, *Tom Bradley's Campaign for Governor: The Dilemma of Race and Political Strategies* 153
SHOSHANA COLMAN

MANUSCRIPT REFEREES, 1988-1989 155

NOTICE TO CONTRIBUTORS 157

CORRECTION

The article "Marginality and Liberalism Among Jewish Elites" by Lerner, Nagai, and Rothman which appeared in the fall 1989 issue of POQ carried a footnote stating that it was "not to be cited or quoted without the permission of the authors." This statement was included by mistake, carried over from the paper when it was still in draft form. As with all POQ articles, citation or quotation (with proper attribution) is welcomed and encouraged.

REFERENDUM DYNAMICS AND THE IRISH DIVORCE AMENDMENT

R. DARCY AND
MICHAEL LAVER

Abstract Referendums are typically interpreted in terms unique to a particular time and place; the 1986 Irish divorce referendum has been no exception. Here this vote is examined in light of referendums in the United States and elsewhere in an attempt to identify more pervasive patterns. It is concluded that there is a referendum dynamic capable of anticipating the diverse fluoridation votes of the 1950s, the ERA campaigns of the 1970s, and the Irish divorce vote of 1986, all of which displayed large voter shifts from an initial majority in favor to an eventual majority against. The dynamic process involved appears to be a pattern of elite retreat brought about by community conflict, rather than the result of confusion, ignorance, a herd instinct, or alienation in the electorate.

Referendums are routinely used in many countries both to amend constitutions and as an alternative to the legislative process. Most referendum outcomes are easily understood in terms of the pattern of opinion in the electorate. Referendum results that are the product of massive short-term shifts in public opinion, the 1986 Irish divorce referendum being a case in point, are difficult to account for in terms of the exogenous preference structure of the electorate, however. Further, such massive opinion reversals are not typical of candidate elections. The fact that opinion reversals do characterize a number of referendum campaigns suggests the operation of some form of short-term "referendum dynamic."

A referendum dynamic would involve a process which begins with

R. DARCY is Professor of Political Science and Statistics at Oklahoma State University and Visiting Professor at the Center for the Study of Irish Elections, University College, Galway. MICHAEL LAVER is Professor of Political Science and Sociology at University College, Galway and was Visiting Professor of Government at Harvard University for 1988–89. John E. Mueller, Richard Engstrom, Hans-Dieter Klingemann, David M. Farrell, and Cornelius O'Leary are appreciated for their help and most useful and critical suggestions.

Public Opinion Quarterly Volume 54:1–20 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

apparent large-scale popular support for some proposal. The proposal is scheduled for a vote, triggering a referendum campaign which ends with a massive short-term opinion reversal and the defeat of the once-popular proposal. The problem is to explain the reversal—something found in many referendum campaigns but virtually absent in candidate elections.

Here an analysis of the Irish divorce referendum explains the shift in voter preferences as the product of an interaction between established elites and fringe activists, an interaction typically absent in candidate elections.

Opinion Reversal: The Irish Divorce Referendum

In April 1986 parties who formed the coalition government in Ireland committed themselves to holding a constitutional referendum to enable the law to be changed to permit divorce. The opposition pledged neutrality and all three national newspapers gave the proposal strong support. The Catholic church, while not permitting Catholics to divorce, did not forbid its communicants from voting in favor, nor did it instruct them how to vote. Catholics were required only to study the teachings of the church before voting according to their consciences (Girvin, 1987; O'Leary and Hesketh, 1987). Public opinion polls had long showed strong popular support for permitting divorce in certain circumstances. This did not translate into a willingness to change the constitution, however, until sometime between late 1985 and early 1986. In February and April 1986, surveys recorded significant majorities in favor of constitutional change (see Table 1).

The vote result in June 1986 was unexpected. On the basis of a 62.1% turnout, divorce was rejected by 36.3% to 63.1% of the popular vote. What is particularly striking about this result is that a short campaign, lasting less than three months, could produce such a massive shift in opinion on an issue that had been discussed and debated for over a decade. Survey work done by the Market Research Bureau of Ireland (MRBI) documents this shift very clearly. MRBI conducted directly comparable polls at the beginning and end of the campaign, as well as a poll in February 1986, before the campaign was mooted, and one in October 1986, long after the furor had died down. The results of these polls, analyzed by major demographic and political groupings, are presented in Table 2.

The most dramatic comparisons are between the February 1986 and the October 1986 polls, the one before the announcement of the referendum, the other well after the vote was recorded. There are no significant differences between them, either overall or in the several catego-

Table 1. Public Opinion on Divorce in Ireland, September 1983–January 1987, MRBI

	Keep Ban on Divorce[a]			Permit in Some Cases[b]			
Date	No	Yes	DK	Yes	No	DK	*N*
September 1983	—	—	—	66	31	3	1,000
January 1984	—	—	—	64	31	5	1,000
May 1984	45	49	6	70	24	6	1,000
October 1984	42	48	10	67	27	6	1,000
February 1985	46	45	9	70	24	6	1,000
April 1985	52	42	6	75	21	5	1,000
July 1985	46	47	8	74	23	3	1,000
November 1985	49	45	6	73	24	3	1,000
February 1986	52	42	6	77	20	3	999
October 1986	51	44	5	70	29	1	1,000

	Change Ban on Divorce			
	For	Against	DK	*N*
April 1986[c]	57	36	7	1,000
June 1986[d]	40	49	11	1,005

SOURCE: Market Research Bureau of Ireland Limited, Dun Laoghaire, County Dublin, Survey Reports 3210/93 (September 1983), 3260/84 (January 1984), 3290/84 (May 1984), 3320/84 (October 1984), 3340/85 (February 1985), 3370/85 (April 1985), 3380/85 (July 1985), 3410/85 (November 1985), 3420/86 (February 1986), 3340/86 (April 1986), 3450/86 (June 1986), 3480/86 (October 1986), all performed for the *Irish Times* and made available through the Center for the Study of Irish Elections, Social Sciences Research Center, University College, Galway.

NOTE: All MRBI polls are based on a multistage clustered probability sample. The selection of either 5 or 10 respondents from randomly selected households within the final cluster is made on the basis of a quota of interlocking demographic controls for age, gender, marital status, and occupation, within area. Interviewing stops when the overall quota of 1,000 respondents has been filled.

[a] "As you probably know, the Irish Constitution contains an article prohibiting divorce. If a referendum was held to remove this article from the Constitution, would you vote to keep the ban on divorce?"

[b] "Do you feel that divorce should be permitted in certain circumstances, or do you feel that it should never be permitted here in Ireland?"

[c] "As you probably know, the Irish Constitution contains an article prohibiting divorce. A referendum is to be held to remove the ban from the Constitution, and in its place provide—through the courts—for the possibility of divorce in the following circumstances:

A marriage has failed.
The period of failure is at least 5 years.
There is no reasonable possibility of a reconciliation.
Dependent spouses and children are provided for.

In the referendum, will you vote for or against this change?"

[d] "As you probably know, a referendum is to be held on Thursday next to remove the ban on divorce from the Constitution, and in its place to provide—through the courts—for the possibility of divorce where a marriage has failed for at least five years, and there is no possibility of a reconciliation. In this referendum, will you vote Yes, for this amendment, or No, against the amendment?"

Table 2. Attitudes Toward Divorce During 1986, by Demographic Group

Group	Remove Ban[a]				Permit in Some Cases[b]		Expect Yes Majority[c]	
	Feb	Apr	Jun	Oct	Feb	Oct	Apr	Jun
Co. Dublin	62	67	45	64	82	72	64	53
Rest of Leinster	46	53	44	44	72	67	61	60
Munster	49	56	32	46	77	70	55	47
Connacht-Ulster	51	52	41	49	75	70	50	50
Urban	59	64	42	59	81	73	64	53
Rural	43	49	38	42	71	66	50	51
Aged 18–24	59	67	49	59	85	83	68	66
25–34	61	64	46	66	84	83	66	59
35–49	55	60	40	54	80	75	60	49
50–64	41	48	33	37	68	53	50	43
65 +	39	39	23	28	58	39	34	42
Nonmanual	60	61	44	60	78	76	60	54
Manual	56	60	40	51	80	69	62	53
Large farmer	29	42	33	41	73	67	51	50
Small farmer	33	50	33	34	61	53	42	46
Male	57	57	49	53	78	70	56	57
Female	48	58	31	49	75	70	60	48
Married	50	57	38	51	77	68	57	51
Single	56	58	44	51	77	74	60	56
Fianna Fail	43	49	30	40	73	62	52	47
Fine Gael	51	66	47	58	81	74	62	56
Progressive Democrat	56	62	68	50	60	81	82	63
Labour	53	54	55	58	72	73	67	63
Other	74	82	60	60	82	69	84	62
DK/refused	56	51	42	56	73	75	55	56
Total	52	57	40	51	77	70	58	52

SOURCE: MRBI reports cited in Table 1.

[a] “As you probably know, the Irish Constitution contains an article prohibiting divorce. If a referendum was held to remove this article from the Constitution, would you vote to remove it and allow for Divorce, or would you vote to keep the ban on divorce?” (February)

“As you probably know, the Irish Constitution contains an article prohibiting divorce. A referendum is to be held to remove the ban from the Constitution, and in its place provide—through the courts—for the possibility of divorce in the following circumstances:

A marriage has failed
The period of failure is at least five years
There is no reasonable possibility of a reconciliation
Dependent spouses and children are provided for.

In the referendum, will you vote for or against this change?” (April)

“As you probably know, a referendum is to be held on Thursday next to remove the ban on divorce from the Constitution, and in its place to provide—through the

Table 2. (Continued)

courts—for the possibility of divorce where a marriage has failed for at least five years, and there is no possibility of a reconciliation. In this referendum, will you vote Yes, for this amendment, or No, against the amendment?" (June)

"As you probably know, the Irish Constitution contains an article prohibiting divorce. If another referendum was held to remove this article from the Constitution, would you vote to remove it and allow for divorce, or would you vote to keep the ban on divorce?" (October)

[b] "Divorce is not legal in Ireland, but a legal separation, where neither party is free to remarry, is permitted. Do you feel that divorce should be permitted in certain circumstances or do you feel that it should never be permitted in Ireland?"

[c] "Irrespective of how you yourself will vote, do you think a majority of voters will vote in favor of this Amendment?"

ries of voters. Comparing them, it is almost as if nothing had happened in between. We find a pattern sustained over most of the period since polling on the issue began: divorce was less favored by older people compared to younger, by rural voters compared to urban, by farmers, and by Fianna Fail supporters compared to the supporters of other political parties.

During the campaign, however, there was a very obvious opinion shift. The June poll, taken just a few days before the referendum, is the only 1986 poll not showing a majority in favor of constitutional change. The clearly "deviant" patterns shown in the June poll suggest strongly that something happened during the campaign which generated a strong but short-term shift in the profile of public opinion on the divorce issue. These polls also contain evidence that the apparent shift in public opinion was not an artifact of differential turnout. There were no significant opinion differences on the issue between those with a high and those with a low likelihood of voting (Jones, 1986b). Comparing the April and June polls in Table 2, we see that the collapse in support for divorce was spread rather evenly across party supporters (with the exception of Labour supporters who showed no shift at all) and also across age groups and social classes. There were, however, some patterns in the opinion shift. It was most pronounced among urban voters and among women.

Opponents of change concentrated their campaign rhetoric very heavily on the plight of married women, who they argued would be cast aside if divorce were to be introduced into Ireland. The possibility that the former breadwinner (always characterized as the man) would remarry, would have legal obligations to support a second family, and would therefore fail to meet maintenance obligations, figured prominently in the campaign of the opponents. The main losers from the introduction of divorce were persistently portrayed as women. The fact that the collapse of support for the amendment was more spectacular among urban voters cannot be explained by a particular argument used in the campaign. Rather, we might look to the generally higher level of electoral volatility of Irish urban rather than rural voters (Laver, 1986a, 1986b; Gallagher, 1987; Sinnott, 1987; Laver, Mair, and Sinnott, 1987). Existing in a much more stable environment, one much more explicitly structured by traditional values, rural voters remained relatively impervious to these changes. They were noticeably less in favor of divorce at the beginning of the campaign, and their views shifted less during it. Particular campaign appeals appear to have been more effective with women and the more susceptible urban electorate. Nonetheless, the relentless movement of opinion in all sectors of the electorate suggests that a more general process was at work, a process that has previously manifested itself in a range of different contexts, ranging from votes over the introduction of fluoride into water supplies to votes to ratify the Equal Rights Amendment (ERA). It is to these that we now turn.

Opinion Reversal: Fluoridation

The Irish divorce referendum was not alone in showing a dramatic opinion reversal. In the 1960s social scientists became curious about fluoridation votes being conducted by a number of American local governments. Public health authorities advocated the addition of small amounts of fluorine to municipal water supplies as a method of reducing tooth decay among children. This roused considerable opposition. Opponents cited potential dangers that included increased susceptibility to cancer, discoloration of teeth, and other problems. The opponents appeared to resent particularly the fact that, if fluorine was added to municipal water, then users of the water would have no option on the matter. They favored alternatives which would permit individuals to make their own decisions concerning use or avoidance of fluorine. The consensus in the scientific community was that putting the fluorine into municipal drinking water was safe, had no adverse side effects, and was the most effective way to control tooth decay among children (see Gamson, 1961; Mueller, 1966; Sapolsky, 1969).

Whereas in the United States most referendums pass (Clubb and Traugott, 1972:141), almost 60% of fluoridation votes failed (Sapolsky, 1969). What is more interesting and potentially relevant to an understanding of the Irish divorce vote, however, was that these fluoridation referendums failed *despite initial overwhelming public support*. Mueller cites one example from Pomona, California. A survey of 532 registered voters in March 1959 found 47% in favor of fluoridation, 19% opposed, and 34% with no opinion, more than two to one in favor. When the actual vote was held in April 1957, however, the voters turned down fluoridation with 36% voting in favor and 64% voting against (Mueller, 1962a, 1962b, 1962c, 1965, 1966, 1968). A similar shift seems to have occurred in other fluoridation votes. Mueller attributes the change to the effective raising of doubts in voters' minds during the campaign, doubts which produced a conservative response as the campaign progressed.

Opponents of fluoridation had several tactical advantages. First, while the bulk of the population was initially favorable, this view was not strongly held nor was it linked to the more central decisions on the part of the voter, such as those relating to political party, ethnicity, religion, or patriotism. Voters could "change their minds without loss of face before others and without running into problems because of some internal need for consistency" (Mueller, 1966:59). Next, the direct benefits of fluoridation were confined to a small group of nonvoters (children) and were largely statistical (and hence, abstract), in that decay would be reduced by a certain probability. On the other hand the dangers (as portrayed by the opponents) were specific and serious (discolored teeth, cancer, bone disease). While the proponents had the support of the bulk of the scientific community, such support was not unanimous. Opponents were able to marshal individuals with medical and scientific credentials as well.

Mueller argues that the opponents needed only to create doubts about fluoridation to win. It was not necessary for the opponents to bring the electorate around to their view (Mueller, 1966:60). The observation that if opponents can create doubts in the minds of voters then voters will respond by voting no has been echoed by a number of authorities since (Sapolsky, 1969; Lee, 1978; Magleby, 1984; Mansbridge, 1986:226). The creation of such doubt during the course of the campaign is not difficult, given the manner in which the debate is typically conducted. By presenting both sides of the argument in an impartial manner, the mass media, voter information sheets, and public meetings can give equal weight to contrasting viewpoints which are not in any way equal in popularity, validity, or reasonableness. A single scientist at variance with his or her discipline, for example, can create the appearance of a deeply divided scientific community when both sides are treated impartially in this way by the media. In the absence of party as a guide to voting in referendums, voters tend to make up their minds late and to be influenced by the positions of prominent individuals (Lee, 1978; Lee, 1981; Magleby, 1984:124). The result is that if opponents are able to counter prominent and expert supporters with prominent and expert opponents, this is usually sufficient to defeat a measure.

Mueller suggests that the fluoridation referendums are unique (Mueller, 1966:67; see also Sapolsky, 1969), and little has been done subsequently to link them with other referendum campaigns. This is a mistake, since there is in fact quite strong evidence to suggest that they are part of a much larger pattern. Magleby, for example, was able to locate 15 California and Massachusetts referendums between 1962 and 1978 in which a series of polls were available (Magleby 1984:211–214). In seven of these there was an opinion reversal from initial support to eventual defeat for the proposition. This initial support ranged from 58% to 79% favorable on issues as varied as control of subversive activities, a lottery, electricity bills, and the control of handguns. The ability of opponents to reverse favorable initial opinion is more general than students of fluoridation campaigns thought. Magleby, in fact, suggests that

> on issues other than those on which voters have standing opinions, there is a predictable movement from general support for the proposition in the early campaign to its rejection as the campaign proceeds. This pattern of opinion change in the direction of the negative vote is an important aspect of many initiative campaigns. Initially, the issue being decided has an appeal that accounts for the early support. As the campaign begins and the voters learn that the proposition has some shortcomings, the appeal of the measure begins to fade. Once the voters begin to learn reasons to vote no, the proportion of negative voters in the cross-sectional surveys begins to increase, and by election day the initiative is often defeated. (Magleby, 1984:167)

Opinion Reversals: State Referendums and the ERA

Jane Mansbridge (1986:14–15) discusses campaigns to ratify equal rights amendments between 1975 and 1986 in New Jersey, New York, Florida, Iowa, Maine, and Vermont. These equal rights amendments typically specified that equality of rights under the law shall not be denied or abridged on account of sex. In each of these campaigns the initial proportion in favor of the amendment exceeded 60%; in each case the campaign resulted in a defeat for the amendment. Most of these states are generally liberal and the political leaders of both major parties (the Republicans and the Democrats) supported ratification, at least in the 1970s (see Boles, 1979; Daniels, Darcy, and Westphal, 1982). Opponents, however, were able to raise what, to voters at least, were serious questions on the consequences of equal rights for women, and the amendments were defeated. Balsam finds a similar shift of 24% away from Welsh devolution during the campaign of 1978–79; again, the result was defeat for what appeared to be a popular referendum proposal (Balsam, 1983).

Bringing the evidence of this extensive series of U.S. state and other referendums to bear on the Irish experience, it would seem that the outcome of the Irish divorce referendum owed less to a set of unique local circumstances and more to some more pervasive referendum dynamic. In the rest of this paper, we consider a number of the processes that various authors have considered might account for this.

The Spiral of Silence

We can deal rather quickly with one possible interpretation of the opinion shift in the Irish divorce referendum, since there is survey evidence that clearly contradicts it. Elisabeth Noelle-Neumann, in her book *Spiral of Silence* (1984) offers a view of public opinion and voting that addresses itself directly to opinion shifts during campaigns. She argues that people, through their general social interaction, develop an intuitive grasp of the relative strength of the various sides of a particular debate. She assumes that, since people want to avoid social isolation, they fail to speak out when they hold the minority view. This generates a tendency for those on the winning side to become more articulate relative to those on the losing side, who tend to become silent. The side that people expect to win thereby appears to be the stronger, and this feeds back into popular perceptions, creating a "spiral of silence" that weakens the "losing" group and strengthens the "winning" group. Popular perceptions can thus become self-fulfilling if the "losers" fear that their cause is hopeless and do not vote. It is of course possible that both sides may initially perceive themselves to be the winner. In this situation, Noelle-Neumann (1984:125) hypothesizes that the side that is the most confident and overestimates its strength the most will be the eventual winner.

The MRBI polls in April and June 1986 enable a direct test of these propositions, since they asked respondents what they thought the outcome of the divorce referendum would be. The results are reported in Table 3. The first hypothesis, that the opinion shift will be in the direction of the outcome predicted by the largest group of voters, gets no support. A clear majority in April and June thought the referendum would pass. The spiral of silence would have favored the pros, not the antis, generating an opinion shift in favor of divorce. What actually happened, of course, was exactly the reverse.

The second Nolle-Neumann hypothesis, that the winner would be the more overconfident, can also be rejected on the basis of these data. In April both groups were equally overconfident, each overestimating their support by 17%. In June, however, the "yes" voters, the eventual losers, were much more (32%) overconfident of their strength than were the antis (2%). A wide margin of the undecided in both polls, furthermore, thought that the measure would pass. All in all, given that

Table 3. Expected Outcome and Intended Vote in the Irish Divorce Referendum

	Intended Vote							
	April 1986				June 1986			
Expected Outcome	Yes	No	DK	Overall	Yes	No	DK	Overall
Majority in favor	74%	33%	55%	58%	72%	37%	50%	52%
Majority against	18	53	22	31	19	51	30	36
No opinion	8	14	23	11	9	12	20	12
Overall	57%	36%	7%	100%	40%	49%	11%	100%
N	570	360	70	1,000	402	492	111	1,005
Overconfidence[a]								
Yes votes		74 − 57 = 17%				72 − 40 = 32%		
No votes		53 − 36 = 17%				51 − 49 = 2%		

SOURCE: Authors' calculations from MRBI reports; see Tables 1 and 2 for wording of questions and source of data.

[a] Expected percent minus intended percent.

the survey data collected during the Irish divorce referendum directly address the concerns of the "spiral of silence" theory, the results for this approach are far from encouraging.

Voter Conservatism

In the Irish divorce referendum, contrary to the predictions of the spiral of silence theory, the opinion shift which took place moved away from the outcome that was expected by the majority of voters. Voters tended to move, furthermore, away from a position that favored change toward a more conservative position. As we have already seen, it has been suggested in relation to other referendum campaigns that there is a tendency for referendum voters to behave in a conservative way. The electorate is thought to be more conservative in referendums, at least in those dealing with social and moral issues, than they are in candidate or party elections (Ranney, 1981:19, 84, 132–133). Magleby finds voters in recent American state referendums have "a tendency to support the status quo by voting no on propositions, and a tendency to vote conservatively on life-style, moral, and racial issues" (Magleby, 1984:191). If this is the case more generally, we should observe major opinion shifts taking place only in one direction, whereby an initial majority in favor of change is reversed as the campaign progresses. On this account, we should not observe opinion shifts away from conservative majorities.

One major case seems to contradict this conclusion, however. This is the *success* of the Italian divorce vote in 1974, in which the electorate on the face of it appears not to have responded in the predicted conservative manner. A closer look at that vote, however, reveals that it conforms more to the "Irish" pattern than appears at first sight. Part of the confusion involves how the vote was reported. Butler and Ranney (1978:230) indicate that 59% voted yes to *repeal* the divorce law. In fact, the question posed was "Do you want the Law of 1 December 1970, No. 898, on the regulation of cases of dissolution of marriage, to be abrogated?" (Clark, Hine, and Irving, 1974:351 n. 55). Thus *a vote of no was a vote for divorce* (the outcome which won) and *a vote of yes (the loser) was a vote against divorce*. The Italian divorce vote had its origins in 1969 when the Chamber of Deputies passed a law (quite similar to that proposed for Ireland) permitting divorce in Italy. Opponents resorted to the initiative, permitted under the Italian constitution, to call a referendum on the issue. While the constitution did permit a popular initiative, the Chamber of Deputies never passed laws to implement this. The result was that the referendum legislation was drawn up *after* the divorce law was passed. For this and for party political reasons, the referendum was considerably delayed; divorce had been part of Italian life for over three years before the referendum was held. This has the very important consequence that the status quo, when the campaign commenced, was the smooth operation of a system of legal divorce. The great unknown was what would happen if divorce was repealed, throwing into question all of the divorces and remarriages that had taken place under it (Beckwith, 1987).

Sten Sparre Nilson suggests that conservative patterns in referendum voting arise because, in the absence of party as a structuring source for opinions, voters fall back on their more conservative instincts (Nilson, 1978; see also Magleby, 1984:174–179). Others suggest that the phenomenon exists because referendums typically deal with a particular issue involving explicit change, and that voters respond by supporting the devil they know rather than the devil they don't know: "no dictator ever had to fiddle a referendum, whether it was Napoleon or Hitler or anybody else; when people are asked to answer a question, one half of which is positive and the other half is negative, and when to give a negative answer looks like peering into the abyss of uncertainty, the people will choose the certainty" (Ronald Butt in Ranney, 1981:13). Stated at this level, however, what is being offered is no more than an assertion based on a pattern of observed behavior. It does not help us to account for why opinion has a tendency to move in a conservative direction during the course of a referendum campaign, though one interpretation is hinted at in Butt's comments about voters' fear of the abyss. It may be that they react against the confusion and uncertainty about the future that the campaign debate inevitably engenders.

Voter Confusion and Alienation

The results of the U.S. fluoridation referendums are often explained in terms of voter confusion or alienation (see Citrin et al., 1975; McDill and Ridley, 1962; Gamson, 1961; Horton and Thompson, 1962; Boskoff and Zeigler, 1964; Aberbach, 1969; but see also Clubb and Traugott, 1972; Sapolsky, 1969; Magleby, 1984). This confusion, it is argued, is produced if the antis mount at least an adequate campaign (see Mueller, 1966; Sapolsky, 1969; Magleby, 1984). The idea is that the antis do not have to convert those who are initially in favor of change; they only need to raise specific doubts about the effects of the proposal or to cause general confusion in the minds of voters. Those confused by the campaign or uncertain about the effects of the proposed change (even if they are predisposed to favor it) will, it is claimed, vote no. This explanation is thus driven by the more or less explicit assumption that voters are risk averse; what happens during the course of the debate that takes place once the campaign gets under way is that all of the potential risks associated with a particular proposed change can form the basis of an effective assault on it by the antis.

Using voter confusion as an explanation implies that voters are not up to the task of deciding complex issues. But confusion is itself a vague concept. None of the existing research measures confusion or even defines the concept in such a way that one could tell if a particular voter is confused or not, and no direct evidence of voter confusion is offered. It is further assumed that confused voters vote no, but there is no evidence in the literature that no voters are more confused than yes voters.

Voter confusion implies a diffuse uncertainty on all aspects of an issue. It might be better to focus on the particulars of what is actually being voted upon. This suggests that there may be a difference between doubts concerning a particular proposal and confusion over a general issue. If opponents successfully suggest that there are potential undesirable defects within a proposal then they have raised doubts. "Confusion" implies that voters are bewildered about what the proposal means—that it is too difficult for them to understand.

Some direct evidence that doubts were raised during the Irish divorce referendum campaign can be found in Table 1. The second column shows the estimated proportion of voters who agreed that "divorce should be permitted in certain circumstances." The first column reports responses to a more direct question about whether people would actually vote to remove the constitutional ban on divorce. The sets of responses differ quite strikingly; many who favor divorce in certain circumstances would not vote to remove the constitutional ban on it, a finding which is a consistent feature of all Irish polls on the subject. It might be felt that this group, 25% of the electorate and 50% of the no vote, was simply confused. However, MRBI asked this particular group of voters, "You say that divorce should be permitted in certain circumstances; why do you also say you would vote to keep the ban on it in the constitution?" The open-ended responses to this question indicated the vast majority of this group

> displayed a notable awareness of the relevant nuances. The general consensus among this group was that while they would favor divorce in certain cases of violence, hardship, child suffering, and where marriage had irretrievably broken down, they are nevertheless afraid of it becoming a free-for-all should the amendment be passed . . . many [did] not see the passing of the amendment as a guarantee that divorce would be introduced only in specific circumstances. (Jones, 1986a:2–3)

Thus, in the actual referendum, voters favoring a limited form of divorce but voting no may not have been confused at all. Rather, they could well have adopted these superficially contradictory positions because they were not certain that only a limited form of divorce could be maintained once the constitution was amended. And we should note that the allegedly inexorable drift that would take place toward "divorce on demand" was another key plank in the campaign of the antis. The amendment introduced by the Irish government attempted to allay these fears by writing restrictions on the scope of divorce directly into the constitution. It addressed the fear of divorce on demand, for example, by requiring that a marriage need to have failed for five years with no possibility of reconciliation. Opponents, however, simply raised other matters that would be made uncertain with the introduction of the amendment, such as the legal status of children, alimony, and inheritance, subjects which the proposal did not deal with at all. The government was eventually forced into a promise to implement legislation giving greater protection to the first family should the referendum pass, a move which served to strengthen the impression that the government proposal was flawed and to increase uncertainty about the real effects of the proposed reform (O'Leary and Hesketh, 1987).

This suggests the shift away from support for the divorce amendment was a shift away from the details of one specific proposal, not a shift away from support for divorce itself. Evidence for this is the October 1986 poll in which the divorce attitudes of the various subgroups of the Irish electorate shift back to the generally prodivorce sentiments found in April before the campaign heated up. Attitudes toward divorce were not changed during the campaign, only attitudes toward the government's divorce amendment.

Overall, the ability of opposition groups to generate voter uncertainty during the course of a campaign, which has been used to explain dramatic opinion shifts elsewhere, appears also to have been a significant factor in the Irish divorce referendum. There is one very important matter that the voter uncertainty explanation does not address itself to, however, which concerns why opinion shifts occur in some referendum campaigns but not in others. Shortly after the divorce referendum there was another constitutional referendum in Ireland that illustrated this point quite clearly.

The Single European Act (SEA) referendum in Ireland was held on 26 May 1987. This had been forced on the government by a finding of the Irish Supreme Court that found the act, which bound the Irish government to "take full account" of the common policies of other European Community member states when forming its own foreign policy, to be in conflict with national sovereignty. It was thus held to be incompatible with Articles 1 and 5 of the Constitution, so that a constitutional amendment would be required before it could pass (Gallagher, 1988, gives a full account of the SEA referendum). The SEA amendment was vigorously opposed throughout the referendum campaign by a number of very active groups. These included parts of the coalition that had successfully opposed the divorce amendment, who argued that the SEA could lead to divorce and abortion being imposed upon Ireland by decisions of the European Court of Justice. Other groups argued that the SEA could reduce Ireland's ability to resist the dumping of nuclear waste in the Irish Sea, and could constrain its independent and more critical foreign policy positions on South Africa and Central America. The opposition, in short, raised many doubts, was very vocal, and received a level of exposure in the Irish media that was more or less equivalent to that of the pro-SEA lobby. Because of this, before the publication of the one and only opinion poll to be conducted during the course of the campaign, there was a popular perception that the amendment was in trouble (Gallagher, 1988:80). In the event, the MRBI poll suggested that only 61% of voters had an opinion on the matter; of these 66% were in favor and 34% against. Twelve days later, on a 44% turnout, the final voting figures were 70% in favor and 30% against.

Thus active, vocal, and well-reported opposition groups attempted to raise a wide range of doubts about the effects of passing the SEA, but they failed signally to generate a conservative opinion shift during the campaign. There is clearly more to generating an opinion shift than placing doubts and uncertainties before the voters. Some other variable is clearly at work, and the Irish examples suggest that this may have something to do with the role played by political elites. In the SEA referendum, there was widespread elite consensus in favor of the amendment, as the four main political parties presented a united front on the issue. Things were quite different in the divorce referendum. Noting the way in which political elites were divided by this bitter and sectarian campaign, as well as in the campaign on the abortion referendum that had shortly preceded it, we turn to an alternative account, based on the notion of community conflict, that may explain why public opinion is more volatile in some campaigns than in others.

Elite Withdrawal and Community Conflict

Between 1950 and 1986, the United States experienced several waves of community conflict over issues that ranged from fluoridation to civil rights to equal rights for women. Often these conflicts became crystalized around a proposal submitted to the city council, the state legislature, or the voters. One interpretation of these conflicts has resulted in the development of what we might think of as the "community conflict" model (see Boles, 1979; Coleman, 1957; Crain, Katz, and Rosenthal, 1969; Mansbridge, 1986). In contrast to more conventional interest group politics, in which conflict is contained by a set of established political institutions and a network of entrenched elites who resolve matters according to well-understood norms and conventions, community conflict develops around sweepingly defined issues with a potential for appealing to broadly based groups within the community at large. Such issues tend to be defined in terms of strongly held abstract symbols, tend to be of fundamental social relevance, and tend either to be nontechnical or to be capable of presentation in a nontechnical form (Cobb and Elder, 1972; Boles, 1979).

Community conflict goes through a particular cycle, according to this account. An established group or coalition of organizations becomes concerned over an issue. After a process of internal dialogue, they negotiate with established political elites, a process which may result in the proposal being submitted by those elites for legislation or public ratification (Boles, 1979:17). Once on the agenda, however, the proposal draws new groups into the debate, often ad hoc groups supported by organizations and networks based outside the established political elite. Participants tend to be of lower social and political status than established political elites and tend to be motivated by very strongly held ideological beliefs, to which the proposed change is seen as a direct challenge. Because these new participants are not part of the traditional political elite (they may even be outsiders to the political system altogether) and because they perceive the situation in terms of threats to their centrally held values, the scope of the conflict is widened into the community at large. Nonelite opponents of change, furthermore, are much less constrained by established political rules and norms (Boles, 1979:17). Wary both of internal splits and of finding themselves on the wrong side of an increasingly divisive issue, established elite organizations begin to develop strategies for avoiding participation. As a result of this, ad hoc groups outside the political elite carry on the struggle for both sides and "finally, with the widening of the conflict, the tone of the debate shifts to one of antagonism, personal slander, and overt hostility" (Boles, 1979:18). Elite withdrawal from an increasingly ugly campaign, fought using techniques that often break the established rules of the game, appears to the public as elite doubt over the issue at stake. The proposed change thereby loses legitimacy and the result is defeat for the proposal.

The community conflict interpretation thus focuses on the withdrawal of established elites as the dynamic element in opinion shifts during controversial political campaigns. These shifts occur when elites either lose or cede control of the campaign to querulous ad hoc groups, a process that is most likely to be set underway during debates over broadly defined nontechnical proposals challenging ideological convictions that are deeply held in important sections of the community as a whole. To what extent can the dramatic opinion shift that took place during the Irish divorce referendum campaign be interpreted in these terms?

Some time before the divorce referendum was announced, a wide network of conservative pressure groups already existed. The most central of these was the Responsible Society, set up in Great Britain in 1971 and in Ireland in 1980, which was active in the successful campaign to introduce a constitutional prohibition on abortion in 1983. This group, with other local conservative Irish groups, provided the basis for the opposition to divorce. Because Fianna Fail had pledged neutrality while the coalition partners were supporting the divorce amendment, opponents of divorce found themselves relying on figures and organizations from outside the political establishment (O'Reilly, 1986). As is typical in examples of community conflict, ad hoc opposition groups in Ireland received literature, funds, and personnel from sympathetic groups located outside the system, in this case from Northern Ireland, Great Britain, and the United States (O'Reilly, 1986; O'Leary and Hesketh, 1987). The result was an emotional campaign in which the focus of the debate shifted away from the specific content of the government proposal as proponents of divorce chained themselves to the gates of the national Parliament while opponents handed out leaflets proclaiming "God Says Vote NO." The campaign became shrill, bitter, sectarian, divisive, and unpleasant (Girvin, 1987; O'Leary and Hesketh, 1987). As was the case with American fluoridation and ERA referendums, well-organized, articulate, and persistent opponents were able to make a real debate in Ireland out of what the government had hoped would be a consensus.

Once the campaign started to get ugly, government ministers began to distance themselves from it. Peter Barry, deputy leader of Fine Gael, foreign minister, and director of the campaign for the leading coalition partner, became "almost apologetic" (*Irish Times*, 30 May 1986:1) over the issue and eventually found occasion to absent himself from Ireland entirely (O'Leary and Hesketh, 1987). Certainly the government coalition, and especially key government members such as Garret FitzGerald and Alan Dukes (Fine Gael) and Dick Spring (Labour), continued their public support of the referendum. Even these big-name politicians, however, found it difficult to gain radio and television exposure, given the absence of "balancing footage" from the other side (*Irish Times*, 14 June 1986:1) or of suitable debate opponents, a situation that was a direct consequence of Fianna Fail's official "neutrality" on the issue. Thus, the activity of these mainstream politicians became marginal to the campaign. Television and newspaper debates were conspicuous by the absence of established political figures, even those still active in the campaign. As soon as most politicians had ducked for cover, those who did not were marginalized.

Overall, some form of explanation along the lines of the community conflict model seems best to fit the facts of the case of the Irish divorce referendum. It suggests that there is indeed a spiral of silence, but that it is a spiral of silence among elites on both sides of the issue, rather than among the voting public at large. As the going gets rough on one side, the elites withdraw, the spiral begins, the going gets rougher, and so on. Once the elites begin to back off, the voters, too, become dubious and vote no.[1]

Conclusion

We may draw several conclusions from this investigation into the Irish divorce referendum. First, the pattern observed in Ireland in the summer of 1986 was far from unique. Instead, the results fit the pattern of most referendums, at least of those dealing with emotional issues. The electorate responded in a conservative manner and the swing from an initial favorable attitude to an eventual negative vote was also something that could be expected, given the pattern of a number of American state and local referendums on fluoridation and equal rights for women.

---

[1] Martin Pugh (1978) makes a similar argument for the *success* of women's suffrage in the United Kingdom in 1918. The war quieted the shrill militant wing of the movement thus moving the question from the streets and newspapers back to traditional parliamentary politics, a move which allowed the constitutional suffragists to succeed with traditional pressure group tactics. Curiously, in 1911 *opponents* of women's suffrage contemplated a referendum on the subject, survey evidence giving them confidence they would win, even among women voters (Pugh, 1978:19–20).

The explanation for this shift from a positive to a negative view is more elusive. Whether or not the community conflict interpretation can be used to account for the dynamics of opinion shifts during particular referendum campaigns, the Irish evidence suggests that it is a useful way to identify those campaigns in which opinion shifts are likely. It suggests that there is indeed a spiral of silence, not among voters (the Irish evidence runs directly counter to that possibility) but among political elites. Situations characterized by community conflict provide the most fertile ground for those who wish to play on voter uncertainties by conjuring up the image of an abyss into which society will fall if the proposal is passed. Opponents of divorce in Ireland seem to have been very successful in developing forceful arguments which raised doubts by playing on voters' fears of change, especially among women, always portrayed by opponents of divorce as the victims of the proposed legislation.

Taken on its own, the account based on "raising doubts" provides an interpretation of the direction of opinion shifts but does not identify campaigns in which these are likely to take place. Taken on its own, the "community conflict" account can identify campaigns in which opinion shifts are likely to take place but does not yet provide an account of the likely direction of these shifts. The evidence of the Irish divorce referendum suggests, however, that these interpretations might productively be blended to provide a fuller account of voter volatility in referendum campaigns.

References

Aberbach, J. (1969)
"Alienation and political behavior." American Political Science Review 63:86–99.
Balsam, Denis (1983)
"Public opinion and Welch devolution." Pp. 197–215 in David Foulkes, J. Barry Jones, and Rick Wilford (eds.), The Welch Veto: The Wales Act of 1978 and the Referendum. Cardiff: University of Wales Press.
Beckwith, Karen (1987)
"Response to feminism in the Italian parliament: Divorce, abortion, and sexual violence legislation." Pp. 153–171 in Mary Fainsod Katzenstein and Carol Mueller (eds.), The Women's Movement in the United States and Western Europe. Philadelphia: Temple University Press.

Boles, Janet (1979)
The Politics of the Equal Rights Amendment. New York: Longman.
Boskoff, A., and Harmon Zeigler (1964)
Voting Patterns in a Local Election. Philadelphia: Lippincott.
Butler, David, and Austin Ranney, eds. (1978)
Referendums: A Comparative Study of Practice and Theory. Washington: American Enterprise Institute.
Citrin, Jack, Herbert McClosky, Merrill Shanks, and Paul Sniderman (1975)
"Personal and political sources of political alienation." British Journal of Political Science 5:1–31.
Clark, Martin, David Hine, and R. E. M. Irving (1974)
"Divorce—Italian style." Parliamentary Affairs 27:333–358.
Clubb, Jerome, and Michael Traugott (1972)
"National patterns of referenda voting: The 1968 election." Pp. 137–169 in Harlan Hahn (ed.), People and Politics in Urban Society. Beverly Hills, CA: Sage.
Cobb, Roger W., and Charles D. Elder (1972)
Participation in American Politics: The Dynamics of Agenda-Building. Boston: Allyn and Bacon.
Coleman, James S. (1957)
Community Conflict. New York: Free Press.
Crain, Robert L., Elihu Katz, and Donald B. Rosenthal (1969)
The Politics of Community Conflict: The Fluoridation Decision. Indianapolis: Bobbs-Merrill.
Daniels, Mark, R. Darcy, and Joseph Westphal (1982)
"The ERA won—At least in the polls." Pp. 15:578–584.
Gallagher, Michael (1987)
"The outcome." Pp. 63–98 in Michael Laver, Peter Mair, and Richard Sinnott (eds.), How Ireland Voted: The General Election of February 1987. Swords: Poolbeg Press.
——— (1988)
"The Single European Act referendum." Irish Political Studies 3:77–82.
Gamson, William (1961)
"The fluoridation dialogue: Is it an ideological conflict?" Public Opinion Quarterly 25:526–537.
Girvin, Brian (1987)
"The divorce referendum in the republic, June, 1986." Irish Political Studies 2:91–96.
Horton, John, and Wayne Thompson (1962)
"Powerlessness and political negativism: A study of defeated local referendums." American Journal of Sociology 67:485–493.
Jones, Jack (1986a)
The Irish Constitutional Referendum, 1986: The Role and Impact of Opinion Polls. Dun Laoghaire: Market Research Bureau of Ireland.
——— (1986b)
"The *Irish Times*/MRBI opinion poll, June, 1986: The constitutional referendum: Article 41." Dun Laoghaire: Market Research Bureau of Ireland.
Laver, Michael (1986a)
"Ireland—Politics with some social bases: An interpretation based on aggregate data." Economic and Social Review 17:107–132.
——— (1986b)
"Ireland—Politics with some social bases: An interpretation based on survey data." Economic and Social Review 17:193–214.
Laver, Michael, Peter Mair, and Richard Sinnott, eds. (1987)
How Ireland Voted: The General Election of February 1987. Swords: Poolbeg Press.
Lee, Eugene (1978)
"California." Pp. 87–122 in David Butler and Austin Ranney (eds.), Referendums: A Comparative Study of Practice and Theory. Washington: American Enterprise Institute.

——— (1981)
"The American experience, 1778–1978." Pp. 46–73 in Austin Ranney (ed.), The Referendum Device. Washington: American Enterprise Institute.
Magleby, David (1984)
Direct Legislation: Voting on Ballot Propositions in the United States. Baltimore: Johns Hopkins University Press.
Mansbridge, Jane (1986)
Why We Lost the ERA. Chicago: University of Chicago Press.
McDill, Edward, and Jeanne Ridley (1962)
"Status, anomia, political alienation, and political participation." American Journal of Sociology 68:205–213.
Mueller, John E. (1962a)
"The politics of fluoridation in Long Beach." Mimeo, Department of Political Science and School of Dentistry, University of California, Los Angeles.
——— (1962b)
"The politics of fluoridation in Los Banos." Mimeo, Department of Political Science and School of Dentistry, University of California, Los Angeles.
——— (1962c)
"The politics of fluoridation in Palo Alto." Mimeo, Department of Political Science and School of Dentistry, University of California, Los Angeles.
——— (1965)
"The politics of fluoridation in Pomona." Revision of 1962 version. Mimeo, Department of Political Science and School of Dentistry, University of California, Los Angeles.
——— (1966)
"The politics of fluoridation in seven California cities." Western Political Quarterly 19:54–68.
——— (1968)
"Fluoridation attitude change." American Journal of Public Health 58:1876–1882.
Nilson, Sten Sparre (1978)
"Scandanavia." Pp. 169–192 in David Butler and Austin Ranney (eds.), Referendums: A Comparative Study of Practice and Theory. Washington: American Enterprise Institute.
Noelle-Neumann, Elisabeth (1984)
The Spiral of Silence. Chicago: University of Chicago Press.
O'Leary, C., and T. Hesketh (1987)
"The Irish abortion and divorce referenda." Paper presented at the annual conference of the Political Studies Association of Ireland, Newcastle, County Down, October.
O'Reilly, Emily (1986)
"The legion of the rearguard." Magill, May, 8–14.
Pugh, Martin (1978)
Electoral Reform in War and Peace, 1906–1918. London: Routledge and Kegan Paul.
Ranney, Austin, ed. (1981)
The Referendum Device. Washington: American Enterprise Institute.
Sapolsky, Harvey (1969)
"The fluoridation controversy: An alternative explanation." Public Opinion Quarterly 33:240–248.
Sinnott, Richard (1987)
"The general election in the republic, February, 1987." Irish Political Studies 2:115–124.

DISARMING THE OPPOSITION
THE PRESIDENT, THE PUBLIC, AND THE INF TREATY

LEE SIGELMAN

Abstract Data from a nationwide opinion survey are used to test the proposition that President Reagan's sponsorship of the INF treaty effectively disarmed the opposition that hard-liners would normally have registered against the treaty. The data are consistent with this interpretation. That is, the president's endorsement of the treaty had only a modest impact on those who perceived the Soviets as relatively nonthreatening, but it had a much more pronounced impact on those who viewed the Soviets as threatening. These results shed new light on the extent to which and the conditions under which Americans are willing to set aside their own policy predilections and follow along behind presidential initiatives in the international arena.

A prominent motif in international politics during the postwar era has been the tendency of national leaders to do exactly what they are expected not to do. It was, for example, Lyndon Johnson, the peace candidate, who presided over the massive American military buildup in Southeast Asia; Charles de Gaulle, the symbol of imperial France, who pulled out of Algeria; Richard Nixon, the consummate Cold Warrior, who opened diplomatic relations with Communist China and initiated détente with the Soviet Union; and Menachem Begin, the former Irgun member, who signed a peace treaty with Egypt.

Developments like these surprise or even shock most observers, who do not expect leaders strongly identified with a particular principle to pursue a policy based on a diametrically opposed principle. However, the recognition is dawning that there are perfectly valid reasons

LEE SIGELMAN is Dean of the Faculty of Social and Behavioral Sciences and Professor of Political Science, Communication, and Public Administration at the University of Arizona. He thanks Gerhard Arminger and Carol Sigelman for their helpful suggestions. The data analyzed in this paper were collected by Market Opinion Research and made available by the Roper Center for Public Opinion Research. Neither Market Opinion Research nor the Roper Center bears any responsibility for the analyses and interpretations presented here.

Public Opinion Quarterly Volume 54:37–47 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

for expecting the unexpected from political leaders. In the words of political scientist Robert Goodin (1983:421), "Politicians' possibilities vary inversely with their declared positions. If an action is somehow out of character for a particular politician, then, for that very reason there are fewer external obstacles to that politician's performing it." The core of this idea is that an action that would otherwise evoke great opposition may well be received favorably if it is undertaken by a leader who would normally be expected to oppose it.[1] As Goodin (1983:421–422) explains:

> we are prepared to accept repugnant policies only from politicians who themselves seem to find them repugnant.
>
> We judge politicians' preferences and political characters from their past records and pronouncements. We rationally come to trust those who seem steadfastly to share our own values and sentiments. If such politicians then suddenly start advocating a policy which they—and we—have always found objectionable, saying that they have reluctantly concluded that it is now necessary or desirable, we would be much more willing to believe this coming from them than had it come from one of our longstanding opponents.

Surely the most notable recent practitioner of the politics of the unexpected has been Ronald Reagan. A longtime advocate of a hardline American posture toward the Soviet Union, which he branded an "evil empire," Reagan returned from his 1985 summit meeting with Soviet leader Gorbachev proclaiming a new spirit of cooperation and trust. This turnabout prompted journalist Sidney Blumenthal (1985:B1) to imagine the hue and cry from conservative Republicans had the president not been one of their own:

> A Democrat, or even a traditional Republican, who returned from a Geneva summit praising Mikhail Gorbachev for being "sincere" would be assailed as a naive appeaser. . . . The Committee on the Present Danger . . . could issue a dire report claiming that the "window of vulnerability" is wide open. And, right-wingers would relentlessly ask: Is it a surprise that such a weak president is willing to give away "Star Wars," our biggest military secret and ultimate defense?
>
> Yet Reagan is unscathed. . . . Though Reagan has said and done things that would be considered weak in a Democratic president, he is seen as strong.

Building on their new spirit of mutual trust, Reagan and Gorbachev held another summit meeting in December 1987, which they concluded by signing a treaty banning intermediate-range nuclear missiles from Europe and the Soviet Union. The signing of the INF treaty provides an excellent vantage point from which to observe how the public responds when a political leader "rises above principle" in the international arena. The question addressed here is whether the sponsorship of the treaty by Ronald Reagan, whose credentials as a hard-liner on foreign and defense policy were well established, in effect "disarmed" the opposition that might otherwise have surfaced against the treaty.[2] The expectations underlying this analysis are straightforward:

---

[1] For supportive evidence based on social psychological experiments, see, e.g., Eagly, Wood, and Chaiken, 1978; for a recent experimental test of this idea focusing on the actions of a hypothetical president in the international arena, see Sigelman and Sigelman, 1986.

1. Within that segment of the American public most favorably disposed toward a cooperative approach to U.S.–Soviet relations, the treaty should have enjoyed widespread support because of its perceived merits, with Reagan's endorsement of it being relatively inconsequential.
2. However, among hard-liners on U.S.–Soviet relations, support for the treaty should have varied as a function of the perceived validation Reagan's active involvement conferred on the treaty.

Data and Methods

The data for this analysis come from a national survey conducted by Market Opinion Research (MOR) between 7 and 14 January 1988, as part of a series of election-year surveys on national security issues involving four polling organizations joined in the "Americans Talk Security" project. Fieldwork for the MOR survey was completed within a month of the signing of the INF treaty.[3] The interviewees were highly supportive of the treaty, as evidenced by their overwhelmingly positive answers when asked: "As you may know, the U.S. and the Soviet Union signed a treaty to eliminate intermediate range nuclear missiles based in Europe and in the Soviet Union. Do you approve or disapprove of this treaty?" Of those expressing an opinion (only 4% did not), 49% indicated strong approval and another 33% said they approved "somewhat." Only 18% were either mildly or strongly opposed to the treaty; analyses not reported here indicated that these opponents were not appreciably different, in terms of partisanship, liberal–conservative sentiments, and other traditional social and political markers, from those who approved of the treaty.[4]

---

[2] It could be argued that President Reagan did not step "out of character" at all, since during his term of office he had indicated on several occasions that he would sign an arms treaty if the Soviets would agree to certain conditions. While the premise of this argument is true, the conclusion does not follow. An out-of-character action need not be a *sudden* deviation from the established norm; it can be a long-standing predilection. The main point is that Reagan had long held that treaties with the Soviets are not worth the paper they are written on and had recently castigated the Soviet Union as an "evil empire." Through these and other actions and public pronouncements, Reagan had left little doubt of his basic distrust of the Soviets. In this sense, his willingness to sign an arms pact with the Soviets clearly ran against the grain of his carefully nurtured public persona.

[3] See MOR (1988) for detailed information about the survey. MOR completed telephone interviews averaging 24 minutes with 1,000 registered voters, who were chosen from households selected on the basis of random digit dialing, with account taken of households having unlisted telephone numbers. The documentation accompanying the data set provides no indication of a response rate for the survey. The sample was stratified according to region, state, and county. Interviewers observed gender quotas and screened respondents for voting status, proceeding only with those who were registered to vote. Age and race discrepancies in the composition of the sample were resolved by weighting the cases without altering the sample size. Reanalysis without weighting of the findings reported here reproduced these findings almost perfectly; the coefficients were either identical or varied by only .01 or .02 in either direction when the weight variable was not invoked.

Each interview also included several questions tapping the degree to which the Soviet Union and its leader were seen as a threat to the United States. On the basis of strong factor-analytic evidence that responses to these items were unidimensional, a composite index was created, with those who perceived the Soviets as highly threatening anchoring one end of the scale and those who perceived little threat from the Soviet Union anchoring the other.[5] This scale was then split at its mean (.57) to differentiate between "low-threat" and "high-threat"

---

4. For example, the rank-order correlations (tau *c*) between support for the treaty and gender (0 = male, 1 = female) and level of education, respectively, are −.10 and .08. Tau *c* is .01 between treaty support and a seven-point party identification scale (1 = strong Democrat, 7 = strong Republican), and −.09 between treaty support and a nine-point political ideology scale (1 = extremely liberal, 9 = extremely conservative). It is thus accurate to say that support for the treaty cut across many of the cleavage lines in Americans' attitudes toward foreign and defense policy.

5. Six items were combined to form the composite scale: (1) "How much of a threat would you say the Soviet Union is to the United States these days: a very serious threat, a serious threat, a minor threat, or not a threat at all?" (2) "Do you think the Soviet Union is a country that we can trust a great deal, somewhat, or not at all?" (3) "If we are weak, the Soviet Union, at the right moment, will attack us or our allies in Europe and Japan" [strongly agree through strongly disagree]. (4) "Is your general impression of Mikhail Gorbachev favorable or unfavorable?" [IF FAVORABLE: "Would that be very favorable or just somewhat favorable?"]. (5) "Do you think Gorbachev is a person we can trust a great deal, somewhat, or not at all?" (6) "Which do you think Gorbachev is probably more interested in—maintaining peaceful relations with the West, or achieving communist domination in the world?" The response categories were rescaled so that each item ranged from 0 to 1, with 1 always indicating a hard-line approach. Principal components analysis yielded one significant factor, which accounted for 49% of the total variance in the six-variable set; all six items loaded strongly and positively on this factor. For each respondent, scores of the six items were then summed and divided by 6 to produce a scale ranging from 0 to 1; coefficient alpha for the scale is .78, indicating acceptable reliability. The scale was dichotomized on the basis of the expectations surrounding the interaction terms. The expectation underlying these terms was that those who perceived the Soviets as threatening would be more influenced by Reagan's endorsement of the treaty than would those who did not view the Soviets as threatening. There was no reason to expect a monotonic relationship to obtain; so there was no reason to think, for example, that if two people both felt highly threatened by the Soviets, the one who felt slightly more threatened would be slightly more influenced by the Reagan validation. That is, the expectation was based on a distinction between fundamental perceptions, not subtle gradations, of threat.

respondents—a crucial distinction, since the two groups are expected to express significantly different opinions of the treaty *and* also to be differentially affected by the "Reagan connection." In point of fact, those who viewed the Soviet Union as highly threatening were less enthusiastic about the treaty than those who took a more favorable view of the Soviets and their leader. Of the latter, 94% expressed support for the treaty, with 68% declaring *strong* support, as compared to 67% and 30%, respectively, of the former.

Responses to another MOR survey question directly tap the significance of the "Reagan connection." After indicating their support for or opposition to the treaty, interviewees were asked to respond to a series of five follow-up questions about the treaty, including the following statement: "President Reagan is well known for his anticommunism, so if he thinks this is a good deal, it must be." Responses to this question enable us to gauge the extent to which certain types of respondents engaged in the expected type of reasoning, i.e., whether those who would not normally be expected to support the treaty were convinced to do so by President Reagan's out-of-character involvement in it. Overall, 61% of the MOR respondents agreed "somewhat" or "strongly" with this statement, including 69% of those in the low-threat category and 54% of their high-threat counterparts.

With these variables established, the expectations outlined above can be stated more precisely as

$$S^* = a + b_1\text{Threat} + b_2\text{Reagan1} + b_3\text{Reagan2} + b_4\text{Reagan3} + b_5(\text{Threat} \times \text{Reagan1}) + b_6(\text{Threat} \times \text{Reagan2}) + b_7(\text{Threat} \times \text{Reagan3}) + e,$$

where:

S^* = the underlying tendency of an individual to disapprove or approve the INF treaty (this tendency is supposed to be metric, but is actually measured on a four-point scale, denoted by S, ranging from strong disapproval through strong approval; using the theory of maximum likelihood estimation, the S^* parameters are estimated in the form of an ordered probit model),

Threat = 0 if the respondent is in the low-threat category, 1 if in the high-threat category,

Reagan1 = 1 if the respondent "disagrees somewhat" on the Reagan validation item, 0 otherwise,

Reagan2 = 1 if the respondent "agrees somewhat" on the Reagan validation item, 0 otherwise,

Reagan3 = 1 if the respondent "agrees strongly" on the Reagan validation item, 0 otherwise,

(Threat × Reagan1) = 1 if the respondent is in the high-threat category *and* "disagrees somewhat" on the Reagan validation item, 0 otherwise,

(Threat × Reagan2) = 1 if the respondent is in the high-threat category *and* "agrees somewhat" on the Reagan validation item, 0 otherwise,

(Threat × Reagan3) = 1 if the respondent is in the high-threat category *and* "agrees strongly" on the Reagan validation item, 0 otherwise,

a, b_1 . . . b_7 = the constant term and maximum likelihood estimates, respectively, in the ordered probit model, and e is the error term.

Responses to the Reagan validation statement are treated as a set of dummy variables due to the ordinality of the underlying scale, with those who expressed strong disagreement serving as the omitted reference category. The model treats support for the treaty as a function of orientations toward the Soviet Union, reactions to the idea that President Reagan validated the treaty by forsaking his established principles, and the interaction between the two. More specifically, b_1 is expected to be strong and negative, reflecting the tendency of unfavorable orientations toward the Soviets to depress support for the treaty; b_2, b_3, and b_4 are expected to be increasingly positive but marginal, reflecting the minor *overall* impact of the Reagan validation factor; b_5 is expected to be negligible, reflecting the unimportance of the Reagan validation factor for those who were already disposed to support the treaty; and b_6 and b_7 are expected to be strong and positive, reflecting the "disarming" effect of Reagan's endorsement on those who would not ordinarily support such a treaty.

Results

The statistical results summarized in Table 1 correspond very closely to these expectations. The performance of the overall model is unexceptional, as evidenced by the highly significant chi square value for the equation, the 54% correct classification percentage, and the .42 rank-order correlation between the actual and predicted values of the dependent variable.[6] More to the point, support for the treaty was

---

[6] Perusal of the entries in the table of expected versus predicted values of the dependent variable reveals that the probit model (a) misclassifies all 76 cases with scores of 0 on the dependent variable (that is, all 76 respondents who strongly disapproved of the treaty) and (b), more generally, does a better job of classifying the respondents who were more supportive of the treaty. These results reflect the logic underlying the classification algorithm. Since evaluations of the treaty were generally positive, it would have taken a stronger combination of factors than were encountered here for the predicted outcome to

Table 1. Support for the INF Treaty: Probit Analysis Summary

A. Summary Statistics

Predictor	MLE	t ratio
Threat	−1.58	−7.58***
Reagan1	−.11	−.47
Reagan2	.10	.51
Reagan3	.50	2.19*
(Threat × Reagan1)	.48	1.78
(Threat × Reagan2)	.77	3.01**
(Threat × Reagan3)	.66	2.35*
Constant	1.90	9.88***
Cutoff points on underlying S^*		
Mu(0)	0	
Mu(1)	0.44	
Mu(2)	1.53	
Chi square/*df*/sig	201.8/7/***	

	Predicted Value on Treaty Support Scale			
Actual Value	0	1	2	3
0	0	35	33	8
1	0	9	33	15
2	0	17	106	105
3	0	15	87	286

Percent correctly predicted	.54
Rank-order correlation (Kendall's tau *b*), predicted vs. actual	.42

B. Projected Probability of Support for the Treaty

Reagan Validation	Low Threat	High Threat
Strongly disagree	.93	.45
Somewhat disagree	.91	.60
Somewhat agree	.94	.77
Strongly agree	.97	.85

C. Projected Probability of Strong Support for the Treaty

Reagan Validation	Low Threat	High Threat
Strongly disagree	.64	.11
Somewhat disagree	.60	.20
Somewhat agree	.68	.37
Strongly agree	.81	.48

NOTES: N = 749. There are 46 missing cases on the measure of support for the treaty, 40 on the Reagan validation, and 190 on the threat variable, which was created from six survey items. With some overlapping missing cases on these variables, this brings the number of cases in the analysis down from 1,000 to 749.

*$p < .05$. **$p < .01$. ***$p < .001$.

strongly influenced by orientations toward the Soviets. Even with the effects of the remaining variables in the model held constant, hard-liners were significantly less enthusiastic about the treaty than were those of a more conciliatory bent. In sections B and C of the table, where the probit coefficients are converted into estimated probabilities of either approval or strong approval of the treaty, we see that no matter how much they agreed or disagreed that Reagan's active involvement constituted a point in favor of the treaty, respondents who rejected the image of the Soviets as untrustworthy and threatening were much more likely to support the treaty than respondents who perceived a Soviet threat. The size of this gap varies from row to row, but it is never less than 12 percentage points and rises as high as 53.

Controlling for the other factors affecting their appraisals of the INF treaty, those who strongly agreed that President Reagan's well-established anticommunist credentials cast the treaty in a favorable light were more likely to back the treaty. The significant effect of the Reagan3 term is represented in the two bottom sections of Table 1 by the tendency of the bottom-row probabilities to exceed those in the other rows. However, this difference is by no means major, and neither the Reagan1 nor the Reagan2 coefficient even approaches statistical significance. Accordingly, it seems fair to characterize the overall impact of the Reagan validation factor as marginal, again consistent with expectations.

This brings us to the heart of the interpretation tested here, the interaction effects. A simple way to convey these effects is to report that among respondents in the low-threat category the correlation (tau *b*) between the Reagan validation variable and support for the treaty was only .13, while among respondents in the high-threat category the correlation was .31. Clearly, then, the Reagan validation factor was more closely related to hard-liners' views of the treaty than to the views of their more conciliatory counterparts. Returning to the probit context, the projections in parts B and C of the table do a good deal to clarify these effects. For those in the low-threat category, these projections confirm that the Reagan validation factor had little or no effect on support for the treaty. This is undoubtedly because no external validation of the treaty was necessary for these respondents, who approved the treaty on its own merits. In sharp contrast, among those in the high-threat category the Reagan validation factor seems to have played a crucial role in shaping evaluations of the treaty. For these

---

have been strong disapproval. Since "somewhat approve" and "strongly approve" (especially the latter) dominated the responses, they also dominated the classification logic; hence, the more enthusiastic the support for the treaty, the greater the probability of being correctly classified.

hard-liners, the probability of expressing support for the treaty would be only .45 if they were unmoved by President Reagan's abandonment of his (and their) normal posture toward the Soviets. However, this probability increases steadily as a function of the Reagan connection, climbing all the way to .85 for hard-liners who took President Reagan's involvement very seriously. As for the probability of expressing *strong* approval of the treaty, the interaction effects point in the very same direction, although the contrasts are less extreme. That is, President Reagan's endorsement of the treaty had a relatively marginal effect on those who would probably have supported the treaty in any event, but effectively disarmed the treaty's most likely opponents.

Conclusion

According to the conventional wisdom of two decades ago, the American public gives the president "a blank check" or "an almost free hand" to do whatever he wants in foreign affairs (Caspary, 1970; Lipset, 1966). Much has changed in recent years with regard to Americans' views on foreign and defense policy and with regard to scholarly understandings of the president's ability to influence these views (see, e.g., Edwards, 1983; Hurwitz and Peffley, 1987; Kernell, 1986). As a consequence of these reconsiderations, the notion that the president holds a blank check now seems little more than a faded memory. Nonetheless, some elements of that idea remain timely, for on several occasions during the 1980s (as in the invasion of Grenada, the abortive bombing of Libya, and the shootdown of an Iranian fighter over the Persian Gulf) the American public has lined up solidly behind the president in times of international strife.

Why does the public so often follow the president's lead in international affairs? When they are called upon to evaluate the president's actions, many Americans respond favorably simply because these actions are congruent with their own preferences. Moreover, irrespective of their substantive preferences, many people habitually fall in behind the president in international situations involving the United States—a tendency John Mueller (1973) dubbed the "rally-around-the-flag" effect. To be sure, an outpouring of public support is less likely when substantive disagreement with the president's actions is widespread (Sigelman, 1980; cf. Kernell, 1984), but for many people the very fact that the president has acted on behalf of the nation is all that matters. To some extent at least, this mentality undoubtedly underlies the support for the INF treaty expressed by many respondents who took a distrustful view of the Soviets *and* strongly rejected the inference that endorsement by a hard-liner like Reagan somehow validated the treaty.

In short, on foreign and defense policy issues Americans may follow the president's lead because they approve of his position on its merits, because they reflexively back the commander-in-chief in foreign and defense policy matters, or for a variety of other reasons, such as a disinclination to second-guess a fait accompli. To this list of reasons should now be added another: even in cases where they would normally respond disapprovingly, Americans may well reconsider their opposition on those dramatic occasions when the president rises above his own deeply held principles. Out-of-character actions disarm the opposition by altering the definition of a situation from one in which the president is seen as acting on the basis of long-standing predilections to one in which his abrupt departure from the well-trod path implies that he must have had unusually compelling reasons for acting as he did.

This is not to suggest that *any* out-of-character action will trigger the same reaction. In an experimental study of a simulated foreign policy crisis, Sigelman and Sigelman (1986) observed that "hawkish" subjects tended to respond favorably when a hypothetical hawkish president acted in a surprisingly conciliatory manner. However, a hypothetical "dovish" president could not get away with stepping out of character in the same simulated crisis; his unexpectedly belligerent actions provoked a bitter backlash among dovish subjects. The case examined here—that of a hard-liner on national security policy negotiating an arms control treaty—closely parallels the one for which Sigelman and Sigelman observed an out-of-character presidential action disarming the opposition. On the basis of this converging evidence, it appears that a certain type of unexpected development—a highly conciliatory move by a president known for long-standing opposition to just such an action—tends to override much of the opposition that would otherwise emerge.

It is far less certain that the opposite type of unexpected development—a highly aggressive move by a president known for a conciliatory approach to international affairs—will have the same effect; one could argue that the American public did follow along behind President Carter's aggressive response to the Afghanistan situation in 1979 and 1980, but it should also be recognized that the president's response was limited in scope, involving no commitment of American military force. Also uncertain is how long the public will tolerate out-of-character actions before it begins to punish its erratic leader for failing to maintain a consistent political persona. Within these limits, however, the very fact that the leader has stepped out of character apparently disarms many of those who would otherwise line up in opposition. This, in any event, seems to be the primary lesson of President Reagan's out-of-character actions surrounding the INF treaty.

References

Blumenthal, Sidney (1985)
"Only Reagan could get away with his nuclear weakness." Washington Post, 8 December, B1–B2.
Caspary, William R. (1970)
"The 'mood theory': A study of public opinion and foreign policy." American Political Science Review 64:536–647.
Eagly, Alice H., W. Wood, and Shelly Chaiken (1978)
"Causal inferences about communicators and their effect on opinion change." Journal of Personality and Social Psychology 36:424–435.
Edwards, George C., III (1983)
The Public Presidency: The Pursuit of Popular Support. New York: St. Martin's.
Goodin, Robert E. (1983)
"Voting through the looking glass." American Political Science Review 77:420–434.
Hurwitz, Jon, and Mark Peffley (1987)
"How are foreign policy attitudes structured? A hierarchical model." American Political Science Review 81:1099–1120.
Kernell, Samuel (1984)
"The presidency and the people: The modern paradox." In Michael Nelson (ed.), The Presidency and the Political System. Washington: CQ Press.
——— (1986)
Going Public: New Strategies of Presidential Leadership. Washington: CQ Press.
Lipset, Seymour Martin (1966)
"The president, the polls, and Vietnam." Transaction, 19–24.
Market Opinion Research (1988)
Americans Talk Security. A Series of Surveys of American Voters: Attitudes Concerning National Security Issues, National Survey No. 2, January 1988, Full Survey Report. Report, codebook, and machine-readable data set circulated by the Roper Center for Public Opinion Research, Storrs, Connecticut.
Mueller, John E. (1973)
War, Presidents, and Public Opinion. New York: John Wiley and Sons.
Sigelman, Lee (1980)
"The commander in chief and the public: Mass response to Johnson's March 31, 1968 bombing halt speech." Journal of Political and Military Sociology 8:1–14.
Sigelman, Lee, and Carol K. Sigelman (1986)
"Shattered expectations: Public responses to 'out-of-character' presidential actions." Political Behavior 8:262–286.

REGULATION VIA THREATS
POLITICS, MEDIA COVERAGE, AND OIL PRICING DECISIONS

STEPHEN ERFLE
HENRY McMILLAN AND
BERNARD GROFMAN

Abstract Using the oil crisis of the late 1970s as a case study, we examine the intertwined influences of public opinion and media attention on the credibility of regulatory threats. We focus on three factors: the intensity of public demands for regulatory intervention, the extent to which there are other competing demands on legislative attention, and the availability of scapegoats external to the industry. We use television news coverage of various topics to measure these three factors. We hypothesize that firms threatened with potential regulation restrained price increases, with the largest and most publicly visible firms exercising the greatest restraint. We find that large, visible oil firms restrained price increases for the most important decontrolled products (diesel fuel oil) when media coverage of the oil industry was extensive. These firms exercised less restraint when the government was busy with other issues or when political instability in the Middle East offered an external rationale for oil price increases.

Under what circumstances will a government intervene to impose emergency price controls on a given commodity? Erfle, McMillan, and

STEPHEN ERFLE is an Assistant Professor of Economics at Dickinson College, Carlisle, PA 17013. HENRY MCMILLAN is a Financial Economist for the Office of Economic Analysis, U.S. Securities and Exchange Commission, Washington, DC 20549. BERNARD GROFMAN is a Professor of Political Science at the University of California, Irvine, CA 92717. The authors are grateful to David Brownstone, Amihai Glazer, and Ken Small for helpful suggestions; to Judith Nerhus and Kevin Easter for able research assistance; and to Wilma Laws and Susan Pursche for final manuscript preparation. The Securities and Exchange Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. The views expressed herein are those of the authors and do not necessarily reflect the views of the Commission or McMillan's colleagues upon the staff of the Commission. The authors are responsible for any errors and omissions.

Public Opinion Quarterly Volume 54:48–63 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

Grofman (1989:136) have proposed the "regulatory threat hypothesis" to explain the response of different firms to the threat of governmental regulation:

> When the threat of governmental price regulation is high, the larger, more visible firms in the threatened industry restrain price increases on those products where price changes are readily apparent to the public; the smaller, less visible firms do not exercise such price moderation. As the threat of governmental price regulation diminishes, firms which had previously exhibited price restraint (and thus whose products are underpriced relative to the market norm) rapidly increase prices to equalize with the industry average.

Our aim in this paper is to explain some of the principal factors that determine when the probability of government price regulation is apt to be high. In particular, we will deal with the joint effects of public and media attention on government regulatory threats. We propose that, other things equal, such intervention will be more likely (1) the more intense the public demand for price regulation, (2) the fewer and less urgent are the competing demands on legislators, and (3) the less available are external scapegoats to whom the industry can attribute responsibility for price increases. We refer to these three factors as the *demand effect*, the *crowding-out effect*, and the *scapegoat effect*, respectively. These three factors affect the credibility of government threats to intervene.

We wish to test these hypotheses using the oil crisis of the late 1970s as a case study. To do so, we first need a measure of the intensity of constituency pressure demanding regulation. Direct measures of public opinion, such as polls and surveys, do not occur with sufficient frequency or regularity to be useful in this context. Therefore, we develop an indirect measure of constituency pressure from television news data. Greater television exposure of a problem signals urgency and may, in a feedback loop, actually increase constituency pressure demanding redress of that problem (and thus probably increase subsequent TV coverage). The demand effect implies that the greater the amount of television time devoted to oil news the greater the regulatory threat. This is consistent with research viewing the media as a major setter of the public agenda.[1]

The government must decide how to react to constituency pressure demanding regulatory intervention. The government's capacity to react changes through time. In particular, there are times when the government is simply too busy with more pressing issues to attend to demands for regulatory intervention. This can be viewed as an information-processing constraint or as a time-and-resource constraint. That is, the government's reaction to constituency pressure depends on factors which limit the government's ability and interest in doing something about that pressure (see Radner and Rothschild, 1975). If the government is busy putting out fires on one front, then it has less time and fewer resources available to attend to other matters. We refer to this as the crowding-out effect.

---

[1] See, for example, Iyengar, Peters, and Kinder, 1982, and Behr and Iyengar, 1985. Parenti (1986:ix) argues that "for many people an issue does not exist until it appears in the news media. . . . Even when we don't believe what the media say, we are still hearing or reading their viewpoints rather than some other. They are still setting the agenda, defining what it is we must believe or disbelieve, accept or reject."

We assume that higher domestic oil prices trigger consumer pressure for oil company regulation. This would be true even if the domestic price increases are caused by international events which affect world oil markets. However, consumer pressure to regulate the domestic industry should be less if the international connection is well publicized than if that connection is not well publicized. Information which places the responsibility for undesirable actions (such as price increases, shortages, or quotas) on some entity other than the domestic industry can reduce the pressure to regulate that industry. We refer to this as the scapegoat effect.[2]

As with the demand effect, we measure the crowding-out and scapegoat effects indirectly using television news data. We describe our empirical measures in the next section. We examine the importance of these three effects in the context of the regulatory threat hypothesis with a simple yet very flexible model of relative firm pricing. In subsequent sections we develop this model and present the empirical results. We examine the period from 1977 through 1980, which includes the high point of the late 1970s oil crisis, the summer of 1979.

Measurement of Constituency Pressure

To determine the importance of the demand, crowding-out, and scapegoat effects vis-à-vis the threat of regulation we must decide on measures for each effect. We measure the demand effect by GOVO, the percentage of news time devoted to federal government activity in the oil industry. This measure does not include general news stories regarding the oil industry, such as exploration activities or the Alaskan pipeline, unless the report discusses the involvement of the federal government.

---

[2] There is similar evidence that presidential popularity does not drop as rapidly in response to negative outcomes when there is a convenient external explanation. See Hibbs, Rivers, and Vasilatos, 1982.

We measure the crowding-out and scapegoat effects with five news variables, all measured in percentage terms, defined as follows:

D: Domestic news regarding the federal government.
This category includes stories about the Supreme Court, HEW, Department of Agriculture, and so on. The only federal government news items excluded from this category are those involving oil issues. (These news items are separately measured by GOVO.)

F: Foreign–U.S. interaction other than Middle East.
This category includes stories, such as SALT and Panama Canal negotiations, which involve interaction of the United States with foreign countries other than those in the Middle East.

ME-US: Middle East–U.S. foreign policy interaction.
This category includes stories, such as the Camp David accords and the Iranian hostage crisis, which directly involve the U.S. government with Middle East governments.

ME-I: Middle East internal political news.
This category includes items, such as hostilities in Lebanon and the Iranian revolution, which do not explicitly involve the U.S. government. It does not include oil market stories, such as those about OPEC.

ME-OIL: Middle East oil market news.
This category includes news items about OPEC and other Middle East oil market stories.

We exclude state and local political news from the domestic category because the federal government wields greater potential for regulatory control of the oil industry. We also exclude stories on natural disasters, human interest, economic events, the weather, and criminal activity unless such stories involve the federal government.

These categories distinguish among political distractions. Categories D, F, and ME-US relate directly to the federal government. These categories may be used to test the crowding-out effect. Because oil industry regulation is a domestic issue, we expect the domestic issues news variable (D) to exhibit the strongest crowding-out effect.[3]

---

[3] There is an issue of how fine a measure to use for the crowding-out effect. This is particularly relevant for the domestic category. Within the federal government, different agencies are responsible for energy regulation and, say, labor market regulation. Energy regulation can proceed while presidential and legislative attention focus on labor market issues. The questions, then, are to what extent legislation can proceed while presidential and legislative attention are directed elsewhere, and to what extent domestic issues deflect presidential and legislative attention from energy issues. At this initial inquiry, we select a relatively broad measure of competing issues.

The federal government's interactions with other nations is separated into two categories due to the unique role that the Middle East plays in the world oil market. United States dealings with non–Middle East nations (F) and Middle East nations (ME-US) may exhibit a crowding-out effect. Their effects should be less than that of domestic issues. ME-US may also exhibit a scapegoat effect.

The final categories (ME-I and ME-OIL) pertain to Middle East internal affairs and international oil market issues that do not directly involve the U.S. government; they provide an external rationale for domestic price increases. Few people would dispute the influence of OPEC on the United States oil market after 1973. OPEC actions provide a scapegoat for domestic oil companies if American consumers believe OPEC decisions caused a domestic oil price increase. In contrast, ME-I provides a broader measure of the scapegoat effect than ME-OIL. Middle East political instability may provide a signal to American consumers of potential oil market instability. Since the public may view Middle East oil market news differently from the rest of Middle East political news, we separately measure these two types of Middle East news.

These categories create a partial partition of the news data; not all news items are included. If a news item is included, it is included in only one of the categories. We placed network news items into categories GOVO, D, F, ME-US, ME-I, and ME-OIL for the period February 1977 through early November 1980 using *Television News Index and Abstracts*. Truncating the sample at the time of the presidential election avoids complications that might arise due to the altered power of President Carter and congressmen who became lame ducks after the 1980 election. The four-year period contained over 3,600 broadcasts. Slightly more than 58% of the approximately 80,000 minutes of network news time was classified into over 21,000 separate entries.[4]

We also noted whether a story was reported within the first five minutes of the broadcast. We constructed a parallel set of variables from this information, denoted with a "5" at the end (e.g., GOVO5). Lead stories may be more accurate than other stories as measures of demand, crowding-out, and scapegoat effects. For example, GOVO5 may better reflect pressure for regulation than GOVO because oil news stories at the ends of broadcasts might be filler and therefore not be as worrisome to firms within the industry.[5]

---

[4] Our coding procedure was as follows. A story was coded only if we judged it to fit one of the categories. The vast majority of stories unambiguously fit or did not fit our categories. If there was a question on whether to include a story, another coder had to agree with the first coder. There were less than 200 questionable stories over the four-year sample period, less than 1% of all coded entries. For each coded story we noted date, network, starting time, stopping time, brief story summary, and category. We performed a series of tests to verify the accuracy of the data. These tests accounted for typing errors and consistency between coded category and story summary.

Empirical Specification

The regulatory threat hypothesis predicts how firms within an industry set prices relative to each other. To test this hypothesis we examine the ratio of product prices charged by two firms rather than the prices of each firm individually. This allows us to reduce the number of descriptive variables to those which vary from period to period rather than examining all variables that affect each firm's pricing decision.

We shall focus on two companies, Asiatic and Exxon. Exxon and Asiatic are among the most important players in the New York harbor market, but they have very different levels of public visibility. Under the regulatory threat hypothesis a highly visible firm, such as Exxon, is more likely to act in an accommodating fashion to the threat of regulatory intervention (i.e., not raise its price) than is a politically invisible but nonetheless important firm, such as Asiatic, acting in the same market. Asiatic is invisible to the public at large because it has no retail operations. One indication of Asiatic's invisibility is the fact that Asiatic was *never* mentioned on network news broadcasts during the 1977–80 period. In contrast, Exxon, the largest domestic major, was mentioned many times.[6]

Let P_a, P_e, and P_s be the logarithms of Asiatic, Exxon, and spot market prices in week t. For notational simplicity, variables do not carry week designations unless they differ from week t; in that case the deviation from week t is noted (for example, -1 refers to the week prior to week t). Define the (logarithm of the) Asiatic-to-Exxon product price ratio in week t to be $AE = P_a - P_e$. Finally, let $dP_i = P_i - P_{i,-1}$ be the percentage change in company i's (or the spot) price during week t, and $dAE = AE - AE_{-1}$ be the percentage change in the Asiatic-to-Exxon price ratio.

The change in the Asiatic-to-Exxon price ratio, dAE, should depend on the price ratio in the previous period relative to its equilibrium level, changes in the spot price, and demand, crowding-out, and scapegoat effects. The relative price adjustment process to be estimated is

$$\begin{aligned} dAE = {} & a_1(AE_{-1} - AE^*) + a_2 dP_s + a_3 Demand \\ & + a_4 Crowding\text{-}out + a_5 Scapegoat + e. \end{aligned} \qquad (1)$$

---

[5] The leading news story in a broadcast is generally the most important story of the day. Foote and Steele (1986) find that two of three networks agree on the lead story 91% of the time and all three agree on the lead story 43% of the time.

[6] Refined petroleum product price data are compiled from *Oil Buyers' Guide*, which has the broadest range of data available for Asiatic and Exxon.

Under stable market conditions, there is some equilibrium value of *AE*, say AE^*. This price ratio is characterized by a stable function of a few explanatory variables, including possible quality differences in the product between companies and differences in the value of having an Exxon versus an Asiatic supply contract. The actual price ratio *AE* need not always be AE^* but will tend toward AE^* if the equilibrium ratio is stable. If the actual price ratio in the previous week exceeds AE^*, we would expect the current price ratio to decline toward AE^*, and vice versa. Put another way, the price adjustment process should be self-correcting; in this case, $a_1 < 0$ describes the rate of disequilibrium adjustment.

The price ratio diverges from AE^* due to short-term disequilibrium effects, including changing price elasticities, production costs, and political factors. We do not need to determine own-price or cross-price elasticities of demand for all firms, however. While these elasticities are needed to discuss an individual firm's price setting they are not needed when examining relative prices across firms. Furthermore, it is unlikely that firm-specific own-price and cross-price elasticities change on a week-to-week basis.

Factor costs to the industry do vary from week to week. Since the spot market allows trade in a nonbranded commodity before its sale as a branded commodity, the spot price represents the marginal cost to each firm. We can therefore use changes in the spot price to measure changes in production costs. Economic theory predicts that a change in the marginal cost common to two firms affects the ratio of the firms' prices according to their relative elasticities of demand; the firm with the lower short-run elasticity of demand is able to pass on those cost increases by increasing its price faster than the firm facing a higher short-run elasticity of demand. In this context, changes in the spot price measure the relative short-run elasticities of demand because short-run elasticities are constant and the spot market equates marginal costs across firms.[7] Therefore, $a_2 > 0$ if Asiatic has a lower short-run elasticity of demand than Exxon and vice versa if $a_2 < 0$.

Regulatory threats may affect the price ratio of two firms which differ in their degree of public visibility. The more visible firm should not increase its product price as much as the less visible firm when a regulatory threat arises. We expect Exxon, the largest domestic major oil company in the United States, to be relatively more visible than Asiatic, an American trading company which is an affiliate of Royal Dutch Shell but which has no retail marketing organization. The higher the demand effect, the more likely is Asiatic's price to be higher than

[7] Furthermore, the Entitlements Program equated the cost of domestic crude across firms. Therefore, the mix of domestic and imported crude oil did not affect marginal costs for any firm (Kalt, 1981).

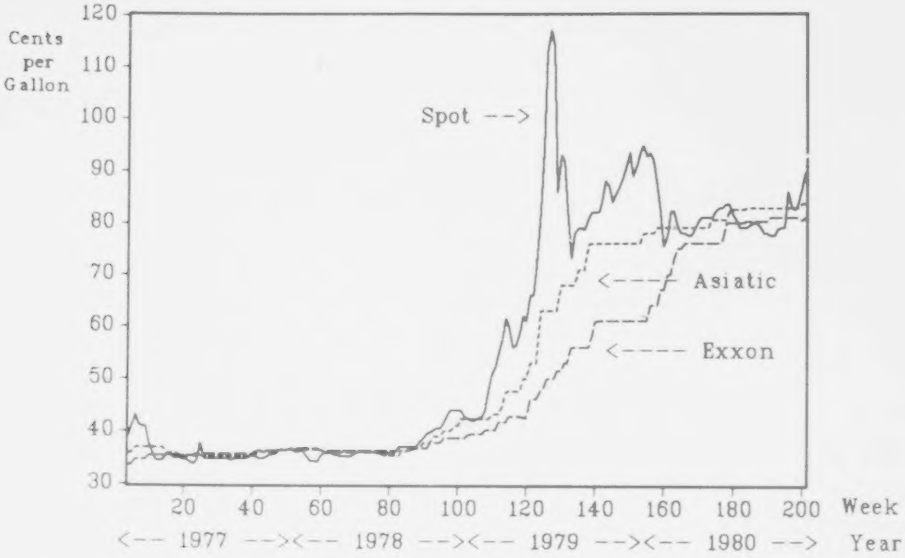


Figure 1. #2 Fuel Oil Prices, 1977–80: Asiatic, Exxon, and Spot

Exxon's, and so $a_3 > 0$. The higher the crowding-out and scapegoat effects, the less likely is Asiatic's price to be higher than Exxon's, and so $a_4 < 0$ and $a_5 < 0$.[8]

In summary, we expect a_1, a_4, and a_5 to be negative and a_3 to be positive; the sign of a_2 depends on the relative price elasticities of Asiatic and Exxon. Since we do not know AE^*, we estimate $a_0 = -a_1AE^*$ instead. GOVO is the primary measure of the demand effect, ME-I and ME-OIL are the primary measures of the scapegoat effect, and D, F, and ME-US are the primary measures of the crowding-out effect.

Empirical Analysis

We estimate equation (1) with weekly data from 1977 through 1980. Excellent weekly wholesale price data are available from *Oil Buyers' Guide* for the most visible decontrolled product, #2 fuel oil (also known as middle distillate, diesel fuel, and home heating oil). Note that gasoline was still subject to price controls in 1979.

Figure 1 shows Asiatic, Exxon, and spot prices for #2 fuel purchased in New York harbor for the sample period. The three prices are essentially the same during 1977 and 1978, as economic theory would predict during stable market conditions. Prices rose in 1979 to a new plateau by the end of 1980. Asiatic let its price follow the spot price more rapidly than did Exxon, in line with the notion that the more visible firm's price lagged the less visible firm's price during crisis periods. Exxon roughly caught up to Asiatic and the spot market by mid 1980.

---

[8] Intuitively, one can think of Asiatic's price as the price that Exxon would charge if no regulatory threat existed. Thus, when $AE > 0$, there is a positive threat of regulation that is keeping Exxon's price below what it would be in the absence of the threat. This interpretation should be tempered by the different adjustment speeds that Asiatic and Exxon may take to fundamental changes in the oil market, as reflected by changes in the spot price.

Table 1 presents summary statistics of regression variables for the full sample period and for 1979 alone. Spot, Asiatic's, and Exxon's prices grew about .4%–.5% per week during the full sample, although most of the increases occurred during 1979. During the crisis year, spot grew at 1.4%, Asiatic at 1.2%, and Exxon at 0.9% per week. The different growth rates reflect Exxon lagging behind Asiatic in increasing its product price. Note further how volatile the spot price was during 1979; the largest one-week spot price increase was 18.4% and the largest one-week spot price decrease was 28.2%. On average during 1979, Asiatic's price was 18.6% higher than Exxon (see *AE* for 1979); the maximum Asiatic-to-Exxon price differential was 30.5%.

Table 1 also shows summary statistics for the news coverage variables. These variables are in percentage terms. Government oil industry news was typically a small fraction of total television news time, but was as high as 16% during one week in 1979. Domestic news stories make up the bulk of news coverage, especially during the 1980 presidential campaign. The ME-US variable is high during 1979 due to the Iranian hostage situation.[9]

Table 2 presents ordinary least squares estimates of equation (1) during the entire four-year Carter presidency and during the crisis year, 1979. These tables lend support to the notions that firms react systematically and differentially to the threat of regulation and that certain media and public opinion factors reduce the threat of regulation. In particular, support is found for the demand, crowding-out, and scapegoat effects.

Before examining the three effects suggested by the regulatory threat hypothesis, we note that the estimated coefficients for the lagged price ratio and the change in the spot price are consistent and robust across all specifications. The constant term is not significantly different from zero for the full sample period, but is for 1979. This signifies that, once other factors are controlled, the firms do not charge different prices in the long run, but did in 1979. The price adjustment process (AE_{-1}) is always significant and in the range of 6% per week during the full

[9] See Erfle, 1988, for a general discussion of the political news categories and how these categories relate to lead news.

Table 1. Summary Statistics

Series	Sample Period: 1977–1980 (Weekly Observations = 198)				Sample Period: 1979 (Weekly Observations = 53)			
	Mean	S.D.	Maximum	Minimum	Mean	S.D.	Maximum	Minimum
dP_a	0.0043	0.0173	0.1728	−0.0275	0.0118	0.0300	0.1728	0.0000
dP_e	0.0045	0.0129	0.0791	−0.0137	0.0091	0.0188	0.0791	−0.0059
dP_s	0.0044	0.0403	0.1841	−0.2819	0.0137	0.0676	0.1841	−0.2819
AE	0.0631	0.0878	0.3054	−0.0237	0.1859	0.0700	0.3054	0.0353
dAE	−0.0002	0.0193	0.1303	−0.0551	0.0027	0.0321	0.1303	−0.0551
GOVO	1.62	2.76	16.36	0.00	4.00	4.07	16.36	0.00
GOVO5	0.72	1.81	13.33	0.00	2.03	2.91	13.33	0.00
D	31.31	11.31	64.57	4.24	24.73	10.16	50.30	4.24
F	9.78	7.57	43.09	0.00	10.08	7.48	31.07	0.57
ME-US	9.06	10.88	53.70	0.00	12.19	16.18	53.70	0.00
ME-I	5.78	6.21	39.92	0.00	6.78	5.49	23.21	0.00
ME-OIL	0.18	0.56	3.28	0.00	0.48	0.88	3.28	0.00

NOTES: dP_a, dP_e, and dP_s are the changes in the logarithm of #2 fuel oil price for Asiatic, Exxon, and spot market.

AE is the logarithm of the ratio of Asiatic to Exxon's price for #2 fuel oil, and dAE is the change in the ratio.

GOVO is government-related oil industry news coverage, and GOVO5 is news coverage beginning in the first five minutes of news broadcast. D is domestic government news, F is foreign news involving the U.S. government but excluding the Middle East, ME-US is Middle East news involving the U.S. government, ME-I is Middle East internal political news, and ME-OIL is Mideast oil market news.

See the text for more precise definitions of the variables. Weekly data, February 1977 through early November 1980.

Table 2. Determinants of #2 Fuel Oil Price Ratios: Factors Affecting the Threat of Regulation

Eq.	Constant	AE_{-1}	dP_s	GOVO	GOVO5	D	ME-I	ME-OIL	$\bar{R}^2$	DW	Sample Period
2.1	.0081	−.063	.122	.0020		−.0002	−.0002	.0009	.150	2.27	1977–80
	(1.66)	(3.7)	(3.8)	(3.7)		(1.84)	(0.99)	(0.38)			
2.2	.0089	−.061	.104		.0034	−.0002	−.0002	.0017	.170	2.27	1977–80
	(1.86)	(3.7)	(3.2)		(4.3)	(1.91)	(1.05)	(0.73)			
2.3	.0666	−.239	.124	.0014		−.0003	−.0028	−.0030	.273	2.11	1979
	(3.25)	(3.5)	(2.08)	(1.36)		(0.76)	(3.04)	(0.68)			
2.4	.0664	−.236	.108		.0026	−.0003	−.0027	−.0024	.293	2.11	1979
	(3.30)	(3.5)	(1.76)		(1.79)	(0.81)	(3.1)	(0.53)			

NOTES: Estimated coefficients of equation (1) in text. Dependent variable in each case is dAE, the percentage change in the ratio of Asiatic to Exxon's price for #2 fuel oil. AE_{-1} is the lagged ratio. dP_s is the percentage change in the spot price for #2 fuel oil. GOVO is government-related oil industry news coverage, and GOVO5 is news coverage beginning in the first five minutes of news broadcast. D is domestic government news, ME-I is Middle East internal political news, and ME-OIL is Mideast oil market news. Weekly data, February 1977 through early November 1980 (N = 198). t statistics in parentheses.

sample period and 24% during 1979. An increase in the spot price raises the Asiatic-to-Exxon price ratio consistently and significantly. This implies that Asiatic had a lower short-run demand elasticity than Exxon and so passed on cost increases more quickly than did Exxon.[10] In the context of long-term supply contracts prevalent in this market, one could say that Exxon was willing to bear more price risk than Asiatic. This suggests more stringent quantity provisions in Exxon's contract than in Asiatic's as a compensating differential.

News stories regarding government interest in the oil industry (GOVO and GOVO5) significantly increased the price charged by Asiatic relative to Exxon for the full sample period. The demand effect is less pronounced in 1979, but GOVO5 remains significant at the 5% level using a one-tail test. While both measures provide similar results, GOVO5 is stronger. This suggests that filler stories are less troublesome to major oil companies than news stories placed earlier in the broadcast. Overall, the results support the demand effect of the theory of regulatory threat.

There is limited support for the crowding-out effect in the estimates. Domestic political news (D) significantly reduces the gap between Asiatic's and Exxon's prices for the full sample period at the 5% level using a one-tail test (equations 2.1 and 2.2). In contrast, D does not exhibit a significant crowding-out effect in the crisis period subsample (equations 2.3 and 2.4).

There is a very strong scapegoat effect during the crisis period, but not during the four-year sample (ME-I is significant in equations 2.3 and 2.4 but not in equations 2.1 and 2.2). Surprisingly, ME-OIL is insignificant; Middle East oil market news does not provide a scapegoat to domestic oil companies. One reason may be that such news is already reflected in spot market prices. Alternatively, such news may truly reflect oil market fundamentals. If so, Exxon would match Asiatic's price increase in this instance. Thus the price ratio *AE* would not change.

Large, visible firms may lag their price adjustments for reasons other than regulatory threats. One reason is that large firms fear the loss of market share. That is, they may expect consumers to retaliate for price hikes by switching to other firms once the market has stabilized itself. Although such behavior has been documented in certain situations (Scherer, 1980:308), it is unlikely to offer an explanation in the current setting. First, consumer retaliation is a more credible concern in a bilateral bargaining situation where the consumers are large or organized. Such is not the case in the #2 fuel oil market, where the consumers are individually small and unorganized. Second, if market share were the primary concern, it is unlikely that firms would also restrict supplies during the crisis. However, Erfle, Pound, and Kalt (1981) document quantity rationing by major oil companies during the 1979 crisis. Therefore, we feel that adverse consequences were more likely to result in the political process than in an unorganized consumer boycott. The political system easily links angry consumers with political entrepreneurs (congressmen) who have the necessary leverage to make consumers' opinions matter.

---

[10] To interpret the error correction mechanism modeled here, suppose that the spot price increased by 10% in week 1 and that the market is modeled by equation 2.1. Then Asiatic and Exxon might both raise their prices, but Asiatic would, on average, increase its price by 1.22% more than Exxon (1.22% = 10 × .122). This would create a gap between the two companies' prices. Asiatic and Exxon would eliminate about 6% of this gap in each subsequent week until their prices were again equal.

We estimated other specifications of equation (1) that are not reported in Table 2. First, we included foreign news (F) and Mideast–U.S. government news (ME-US) in other specifications. These variables were never significant. Second, we substituted other prominent American oil companies for Exxon. Similar results obtain for Gulf, Mobil, Texaco, and Citgo. The effect of regulatory threat is stronger the larger the domestic major.[11] Third, we tested whether the results differed when D5 replaced D, when F5 replaced F, and so on. The results did not qualitatively differ from those in Table 2. Unlike GOVO the full news report versions of each distraction variable provided a better fit and higher t statistics than did the first-five-minute versions of each distraction variable. Fourth, we tested and found no evidence of simultaneous equation bias. Erfle and McMillan (1989) find that television news coverage of the oil industry depends on spot price changes and Asiatic price changes. However, neither the level nor the change in the Asiatic-to-Exxon price ratio affects news coverage. So while television merely reports general oil industry developments, television news reports influence the product prices that politically sensitive firms charge.

Conclusions

We argue that three factors affect the threat of regulation: the intensity of public opinion demanding intervention, the government's capacity to react to that pressure, and the availability of external scapegoats which deflect public attention from regulation. We find empirical support for each of these contentions. Focused public opinion, as measured by television coverage of the oil industry, differentially affected oil firms' product prices during the 1977–80 period. Measures of the government's ability to react to that pressure also affected relative prices for petroleum products.

---

[11] All of these companies are more visible than Asiatic. Exxon is the largest and, arguably, the most visible domestic major. Erfle, McMillan, and Grofman (1989) examine evidence of the demand effect across these companies and find that the speed of price adjustment is inversely related to firm size.

We propose that highly visible firms moderate price increases relative to less visible firms when the threat of regulatory intervention is deemed to be serious. Our measure of constituency pressure demanding regulatory redress (derived from television news coverage of the oil industry) was a highly significant predictor of relative pricing behavior. The more visible firm (Exxon) moderated price increases relative to the less visible firm (Asiatic) when faced with threats of regulatory intervention. Our work also suggests the possibility that "jawboning" strategies, including the judicious use of threats and public exposure, may alter the behavior of the more politically sensitive firms, at least during market crises. However, the evidence suggests that jawboning strategies can work on a temporary basis only.[12]

We also identify two effects which may reduce the threat of regulation: the crowding-out effect and the scapegoat effect. The crowding-out effect pertains to situations where other political issues crowd out demands for regulation from the political agenda; the scapegoat effect pertains to situations where price increases can be attributed to some entity other than the domestic industry (e.g., the Middle East), thereby reducing the pressure to regulate the domestic industry. We find empirical support for both effects: the crowding-out effect was strongest over the entire 1977–80 time period, and the scapegoat effect was strongest during the oil crisis in 1979.

The results presented in this paper demonstrate that highly visible firms react to a threat of regulatory intervention as signaled by media attention focused on the industry. Our paper thus makes a contribution to the theory of regulation, which has hitherto almost exclusively focused on regulatory efficiency and regulatory compliance (Stigler, 1971; Wilson, 1980) or on the power of firms to affect government action (Salamon and Siegfried, 1977) and not on the role of regulatory threat in affecting decisions of the firm (Hunt, 1974; Erfle, Pound, and Kalt, 1981; Block, Nold, and Sidak, 1981; Olmstead and Rhode, 1985), as well as to the theory of media effects.

---

[12] Whether the effect of media and regulatory threat on firm behavior is good or bad depends on normative criteria. The evidence presented here suggests that political pressure transmitted via the media may have contributed to petroleum shortages during 1979. Within the standard economic analysis this implies allocative inefficiency. However, if equity was also a goal of energy regulation at this time, as Harrison (1981) and Kalt (1981) argue, regulatory threats may well have contributed to achieving that goal. Relative weights for these two considerations is beyond the scope of this study and subject to further research.

References

Behr, Roy, and Shanto Iyengar (1985)
"Television news, real-world cues, and changes in the public agenda." Public Opinion Quarterly 49:38–57.

Block, Michael K., Frederick Carl Nold, and Joseph Gregory Sidak (1981)
"The deterrent effect of antitrust enforcement." Journal of Political Economy 89:429–445.

Erfle, Stephen (1988)
"Network news coverage of political and lead news stories during the Carter presidency." Mimeo, University of California, Irvine.

Erfle, Stephen, and Henry McMillan (1989)
"Determinants of network news coverage of the oil industry during the late 1970s." Journalism Quarterly 66:122–129.

Erfle, Stephen, Henry McMillan, and Bernard Grofman (1989)
"Testing the regulatory threat hypothesis: Media coverage of the energy crisis and petroleum pricing in the late 1970s." American Politics Quarterly 17:132–152.

Erfle, Stephen, John Pound, and Joseph Kalt (1981)
"The use of political pressure as a policy tool during the 1979 oil supply crisis." Discussion Paper no. E80-09. Kennedy School of Public Policy, Harvard University.

Foote, Joe S., and Michael E. Steele (1986)
"Degree of conformity in lead stories in early evening network TV newscasts." Journalism Quarterly 63:19–23.

Harrison, David (1981)
"Regulation and distribution." Pp. 185–208 in Allen R. Ferguson (ed.), Attacking Regulatory Problems: An Agenda for Research in the 1980s. Cambridge, MA: Ballinger.

Hibbs, Douglas A., R. Douglas Rivers, and Nicholas Vasilatos (1982)
"On the demand for economic outcomes: Macroeconomic performance and mass political support in the United States, Great Britain, and Germany." Journal of Politics 44:426–462.

Hunt, Michael S. (1974)
"Trade associations and self-regulation: Major home appliances." Pp. 39–55 in Richard E. Caves and Marc J. Roberts (eds.), Regulating the Product: Quality and Variety. Cambridge, MA: Ballinger.

Iyengar, Shanto, Mark Peters, and Donald Kinder (1982)
"Experimental demonstrations of the 'not-so-minimal' consequences of television news programs." American Political Science Review 76:848–858.

Kalt, Joseph P. (1981)
The Economics and Politics of Oil Price Regulation. Cambridge, MA: MIT Press.

Oil Buyers' Guide (1977–80)
Weekly issues. Lakewood, NJ: Petroleum Publications, Inc.

Olmstead, Alan, and Paul Rhode (1985)
"Rationing without government: The west coast gas famine of 1920." American Economic Review 75:1044–1056.

Parenti, Michael (1986)
Inventing Reality: The Politics of the Mass Media. New York: St. Martin's Press.

Radner, Roy, and Michael Rothschild (1975)
"On the allocation of effort." Journal of Economic Theory 10:358–376.

Salamon, Lester, and John Siegfried (1977)
"Economic power and political influence: The impact of structure on public policy." American Political Science Review 71:1026–1043.

Scherer, F. M. (1980)
Industrial Market Structure and Economic Performance. 2d ed. Chicago: Rand McNally College Publishing Company.

Stigler, George (1971)
"The theory of economic regulation." Bell Journal of Economics and Management Science 2:3–21.
Television News Index and Abstracts (1977–80)
Monthly issues. Nashville, TN: Vanderbilt Television News Archive.
Wilson, James (1980)
"The politics of regulation." Pp. 357–394 in James Q. Wilson (ed.), The Politics of Regulation. New York: Basic Books.

THE SOCIAL CONTEXT OF VOTER TURNOUT

BRUCE C. STRAITS

Abstract This paper reexamines the classic interpersonal explanation of electoral participation. The analysis is based on loglinear modeling of published tabulations from the November 1984, 1982, and 1980 Current Population Surveys. The results support a contagion model of spousal voting. Turnout is higher among those living with a spouse than among other members of the electorate. The high association between husband's and wife's turnout noted in a 1956 presidential election survey ($Q =$ .94) is found to be robust and stable over time. The relevance of these findings for contemporary models of electoral participation is discussed.

The role of personal relationships in turning out voters is absent or negligible in contemporary models of electoral participation (Kinder and Sears, 1985; Aldrich and Simon, 1986; Grofman, 1987).[1] In contrast, the classic Columbia voting studies (Lazarsfeld, Berelson, and Gaudet, 1948; Berelson, Lazarsfeld, and McPhee, 1954) delineated systematically the social context of voting behavior. Primary-group associates (friends and family) were theorized as providers of political information and influence.

The likely mechanisms by which social environments stimulate turnout include the channeling of political information, persuasive argumentation, social pressure, cooperation, and friendly competition. A voter's family, friends, and coworkers serve as bridges to larger networks of social relations (Berelson, Lazarsfeld, and McPhee, 1954:94) which provide access to the information and resources that stimulate political interest and provide an informed basis for making ballot choices. Interpersonal influence occurs when the reluctant voter adopts or acquiesces to an associate's expectations, when associates

---

BRUCE C. STRAITS is Associate Professor of Sociology at the University of California, Santa Barbara. The author thanks Noah E. Friedkin, Eric R. A. N. Smith, and Thomas P. Wilson for comments on an earlier draft.

1. A notable exception is the network approach of Huckfeldt and Sprague, 1987, 1988.

Public Opinion Quarterly Volume 54:64–73 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

reciprocally support each other's electoral participation, and even when an individual is motivated to vote for the opposing candidate to cancel out an associate's choice.

This paper provides evidence that one crucial aspect of social context, the spousal relationship, cannot be neglected in explanations of electoral participation. The results support a contagion model of voting: turnout is higher among those living with a spouse, and spouses tend to be joint voters or nonvoters. The similarity in spousal voting appears quite stable over time and unaffected when controls are applied for the major nonattitudinal determinants of voter turnout (educational attainment and age).

Relevant Research

Thirty-one years ago in this journal William A. Glaser (1959) chided political researchers for ignoring social relationships by studying electoral participation as an individual rather than a collective process. Examining turnout among married couples in the 1956 presidential election, Glaser (1959:569) concluded that voting "tends to be a joint household activity, with the members either voting or staying home as a unit." This paper underscores Glaser's conclusion by demonstrating that the high association he found between husband's and wife's turnout ($Q = .94$) is robust and stable over time.

Although other early studies dealt largely with the similarity of respondents' and associates' political views and candidate preferences, passing reference was made to turnout congruence among family members and among friends (Lazarsfeld, Berelson, and Gaudet, 1948:140–145; Berelson, Lazarsfeld, and McPhee, 1954:93–101, 337; Campbell, Gurin, and Miller, 1954:202). Other research supporting interpersonal influence includes evidence that preelection contacts by party representatives or by survey interviewers increases turnout (for a summary of these studies, see Kinder and Sears, 1985:702–703; Berelson, Lazarsfeld, and McPhee, 1954:337).

Data and Modeling

In November of presidential and congressional election years, persons of voting age in Current Population Survey (CPS) sampled households are asked, "In any election some people are not able to vote because they are sick or busy or have some other reason, and others do not want to vote. Did [you/another household member] vote in the election on November 6th [1984]?"[2] The published CPS information on spousal turnout is limited to three-variable tables displaying husband's and wife's turnout by a single characteristic of the husband or of the family. Log-linear modeling will be used primarily to test if the association of husband's and wife's turnout varies by husband's age or education.

The published CPS voting figures are derived by inflating weighted sample results to match independent population estimates. For this study, these (inflated) numbers were downwardly adjusted prior to log-linear analysis to compensate roughly for sample design effects.[3] Although the downward adjustment is far from a perfect correction for the complexities of the CPS estimates, it permits the use of significance tests as a practical criterion for screening out noise and trivialities in the data analysis.

Results

Evidence from early studies suggests that political homogeneity among close associates increases over time. Berelson, Lazarsfeld, and McPhee (1954:94–98) report that similarity of political party preference with friends is greater for older than for younger respondents. In *The People's Choice* panel study of the 1940 presidential election, family members became more homogeneous as the campaign continued (Lazarsfeld, Berelson, and Gaudet, 1948:144).

To explore the possibility that spousal turnout similarity increases with age, the log-linear models displayed in Table 1 were fitted to the November 1984 CPS tabulation of wife's turnout (W) by husband's turnout (H) by husband's age (A). Model 6 implies two-way associations between husband's and wife's turnout, {HW}; wife's turnout and husband's age, {WA}; and husband's turnout and husband's age, {HA}. This model fits the data quite well; it yields an insignificant likelihood-ratio chi-square test statistic, $L^2 = 4.5$ with 4 degrees of freedom (*df*),

---

[2] Questions of this type produce slightly inflated estimates of turnout (Silver, Abramson, and Anderson, 1986; U.S. Bureau of the Census, 1986:9–10). A more serious validity issue for this study is the use of proxy respondents to report on the voting of others in the household; approximately 40% of the 1984 CPS voting results are based on proxy respondents (U.S. Bureau of the Census, 1986: table 18). Although a special 1984 verification test disclosed that proxy and self-reports were in agreement about 99% of the time (U.S. Bureau of the Census, 1986:10), the proxy reporting is likely to positively bias reported spousal voting homogeneity.

[3] For example, each cell count (population estimate in thousands) in the 1984 tables was divided by 3.223 to deflate it to the number of sample cases that would produce the same sampling error as the CPS (U.S. Bureau of the Census, 1986: appendix C) had simple random sampling been employed. This approach is used commonly to adjust published CPS data for sample design effects (Hauser and Featherman, 1977: appendix F; Duncan and Stenbeck, 1988:4).

Table 1. Log-Linear Models of Husband's and Wife's Voting Turnout by Husband's Age, 1984 Presidential Election

Model	*df*	L^2	L^2/df	*p*
1. {W} {H} {A}	13	6392.0	491.7	<.001
2. {WA} {H}	9	5802.8	644.8	<.001
3. {HA} {W}	9	5747.6	638.6	<.001
4. {HA} {WA}	5	5158.5	1031.7	<.001
5. {HA} {HW}	8	128.6	16.1	<.001
6. {HA} {WA} {HW}	4	4.5	1.1	.343
7. {HWA}	0	0.0		

Fitted Model 6

	Odds Husband Voted			
Husband's Age	Wife Voted	Wife Did Not Vote	Odds Ratio	*Q*
18–24 years	4.949	0.202	24.4	.92
25–34 years	6.263	0.256	24.4	.92
35–44 years	7.865	0.322	24.4	.92
45–64 years	11.953	0.489	24.4	.92
65 years and over	13.912	0.569	24.4	.92

SOURCE: U.S. Bureau of the Census, 1986: table 17.

NOTES: H = Husband's turnout (vote/not), W = Wife's turnout (vote/not), A = Husband's age

indicating that the fitted and observed cell frequencies are not significantly different. Comparison of models 6 and 7 confirms that interaction among the three variables, {HWA}, is nonsignificant ($p = .343$). Thus, it is reasonable to conclude that spousal turnout homogeneity is not conditional on husband's age.

The second panel of Table 1 displays summary measures of association based on the *fitted* frequencies of model 6. The first two columns show the odds that a husband voted conditional on the husband's age and the wife's turnout.[4] For example, the odds of voting for a 45–64-year-old husband whose wife voted is 11.953 (almost 12 to 1); if the wife didn't vote, the conditional odds falls to 0.489. The association between husband's and wife's turnout is summarized in column 3 with odds ratios, computed by dividing the conditional odds in column 1 (voting wife) by column 2 (nonvoting wife). These odds ratios, which are identical since spousal turnout association is independent of husband's age in model 6, indicate that the odds of a husband voting is about 24.4 times greater if the wife voted than if she did not vote. Column 4 displays the odds ratios converted to values of Yule's *Q*, a normed measure of association.

---

[4] Turnout odds are shown for the husband since the published CPS tables display husband's but not wife's characteristics. The odds ratios, however, are the same whether computed from husband or wife turnout odds.

The spousal turnout association estimate ($Q = .92$) is quite close to that ($Q = .94$) found in the 1956 presidential election survey reported by Glaser (1959:564–565). The age voting pattern for husbands also is consistent with previous studies: turnout rises with increasing age and may decline slightly at advanced ages (Wolfinger and Rosenstone, 1980; Jennings and Markus, 1988; Duncan and Stenbeck, 1988:10).

Educational attainment is another causally relevant factor available in the CPS tables. Indeed, Wolfinger and Rosenstone (1980:102) found formal education to be the most important variable in explaining differences in turnout:

> The personal qualities that raise the probability of voting are the skills that make learning about politics easier and more gratifying and reduce the difficulties of voting. Education increases one's capacity for understanding complex and intangible subjects such as politics, as well as encouraging the ethic of civil responsibility. Moreover, schools provide experience with a variety of bureaucratic problems, such as coping with requirements, filling out forms, and meeting deadlines.

Furthermore, these relevant skills and experiences with politics may be acquired outside of school through accumulated life experiences, which may explain why turnout increases with age (Wolfinger and Rosenstone, 1980:58–60) even though older Americans have less formal education.

Wolfinger and Rosenstone's thesis is extended easily to spousal turnout: the schooling and relevant life experiences of each spouse may mutually influence the other's turnout. The political interests and skills of a better-educated spouse, for example, may be sufficient to turn out an otherwise disinterested partner. The published CPS spousal turnout tabulations, unfortunately, show only husband's educational attainment. Table 2 displays log-linear modeling of husband's turnout (H) by wife's turnout (W) by husband's education (E) for the 1980 presidential election.[5] Spousal turnout homogeneity varies slightly by husband's

---

[5] The 1980 CPS was used instead of the 1984 study (U.S. Bureau of the Census, 1986: table 17) because the latter's marginal education distribution is in error (e.g., too few college graduates). Nevertheless, a log-linear analysis of the 1984 data (with a collapsed category of four years or more of college) produced results similar to those reported here for 1980.

Table 2. Log-Linear Models of Husband's and Wife's Voting Turnout by Husband's Education, 1980 Presidential Election

Model	*df*	L^2	L^2/df	*p*
1. {W} {H} {E}	22	8715.9	396.2	<.001
2. {WE} {H}	15	7670.7	511.4	<.001
3. {HE} {W}	15	7571.2	504.7	<.001
4. {HE} {WE}	8	6526.0	815.8	<.001
5. {HE} {HW}	14	226.2	16.2	<.001
6. {HE} {WE} {HW}	7	27.3	3.9	<.001
7. {HWE}	0	0.0		

Model 7 Frequencies

	Odds Husband Voted			
Husband's Education	Wife Voted	Wife Did Not Vote	Odds Ratio	*Q*
0–4 years	3.702	0.241	15.4	.88
5–7 years	4.976	0.363	13.7	.86
8 years	6.247	0.298	20.9	.91
9–11 years	6.373	0.238	26.8	.93
12 years	9.013	0.319	28.2	.93
13–15 years	11.889	0.459	25.9	.93
16 years	19.951	0.700	28.5	.93
17 years or more	29.071	0.678	42.9	.95

SOURCE: U.S. Bureau of the Census, 1982: table 9.
NOTES: H = Husband's turnout (vote/not), W = Wife's turnout (vote/not), E = Husband's education.

education (model 7), as adding the three-way interaction, {HWE}, to model 6 significantly improves fit (L^2 reduction of 27.3, 7 *df*, $p < .001$).[6] However, since substantively inconsequential differences are likely to be statistically significant in large samples, the simpler model 6 also provides a reasonable fit to the data ($L^2/df = 3.9$).

If one ignores the slight three-way interaction in favor of fitting model 6, the estimated spousal turnout association in 1980 (odds ratio = 25.7, Yule's $Q = .925$) is quite close to the estimates for the 1956 (Glaser, 1959) and 1984 (Table 1) presidential elections. Consequently, spousal turnout homogeneity appears quite stable over time and unaffected when controls are applied for the major nonattitudinal determinants of voter turnout (educational attainment and age).

---

[6] The strength of husband's and wife's turnout association increases very slightly with husband's education, an interaction which could reflect just the absence of age controls, unresolved sampling complexities, or voting overreporting errors linked either to education (Silver, Abramson, and Anderson, 1986) or to the sex of proxy respondents.

Finally, if voting is a contagious process, turnout should be higher among members of married-couple than single-person households (Wolfinger and Rosenstone, 1980:44–46). The influence of a spouse or other intimate associate may be sufficient to prod an otherwise unmotivated individual to vote. If spouses mutually influence and amplify each other's voting dispositions, slight spousal differences in background and political attitudes may produce large effects on turnout.

Table 3 displays log-linear models of reported voter turnout (V) in the 1982 congressional election by whether a spouse is present (P) and by age (A), estimated separately for males and females. Variable *P* is a dichotomy comparing married persons living with a spouse with all others (single, widowed, divorced, and married persons living apart). Only the saturated model (5) fits the data for both sexes. It implies that both age and the presence of a spouse influence turnout, and the effect of having a spouse present (or age) on turnout varies by age (or spouse presence). Voting odds based on model 5 confirm that, except for 18–24-years-olds, turnout is higher among those living with a spouse.[7]

Discussion and Summary

To some extent, similarities in spousal turnout may reflect self-selection of like-minded mates and shared family circumstances, such as mutual nonvoting due to transportation problems, illnesses, or length-of-residence registration requirements. The limitations of the published CPS voting tabulations preclude definite testing of self-selection or other alternative explanations. At best, controlling on husband's age or education, factors which may operate to inflate spousal homogeneity spuriously by means of spouse selection and proxy reporting errors, did not affect spousal voting congruence appreciably. Also, separate controls for shared household characteristics, family income, or duration of residence fail to affect the observed association of husband's and wife's turnout (U.S. Bureau of the Census, 1986: table 17; 1982: table 9).

Instead of representing rival hypotheses to be ruled out, self-selection and shared circumstances should be incorporated into an explication of voter turnout. That is, within the constraints imposed by their social environments individuals construct their interactional net-

---

[7] The nonvoter denominator of the turnout odds includes citizens reported as "did not vote," "do not know," and "not reported" (U.S. Bureau of the Census, 1983:39).

Table 3. Log-Linear Models of Voter Turnout in the 1982 Congressional Election by Presence of a Spouse, Age, and Sex

Model	*df*	L^2	L^2/df	*p*
Males				
1. {PA} {V}	9	2913.7	323.7	<.001
2. {PA} {VA}	5	457.6	91.5	<.001
3. {PA} {VP}	8	1531.2	191.4	<.001
4. {PA} {VA} {VP}	4	115.6	28.9	<.001
5. {PAV}	0	0.0		
Females				
1. {PA} {V}	9	2769.2	307.7	<.001
2. {PA} {VA}	5	424.8	85.0	<.001
3. {PA} {VP}	8	2034.1	254.3	<.001
4. {PA} {VA} {VP}	4	109.9	27.5	<.001
5. {PAV}	0	0.0		

Model 5 Frequencies

	Odds Voted			
Age	Spouse Present	Others	Odds Ratio	*Q*
Males				
18–24 years	0.318	0.369	0.86	−.07
25–44 years	1.031	0.629	1.64	.24
45–64 years	2.117	0.958	2.21	.38
65–74 years	2.757	1.060	2.60	.44
75 years and over	2.022	0.939	2.15	.37
Females				
18–24 years	0.305	0.365	0.83	−.09
25–44 years	1.111	0.689	1.61	.23
45–64 years	2.123	1.130	1.88	.31
65–74 years	2.244	1.425	1.58	.22
75 years and over	1.517	0.793	1.91	.31

SOURCE: U.S. Bureau of the Census, 1983: table 7 (U.S. citizens).
NOTES: V = Turnout (vote/not), P = Spouse present (yes/no), A = Age.

works (Huckfeldt and Sprague, 1987, 1988), which in turn confirm and amplify (positive feedback) prior voting or nonvoting dispositions.

Contemporary models of electoral participation treat individuals in isolation from their social environment and primary network connections; they predict the same turnout for individuals with the same independent-variable values regardless of the social context within which voting takes place. However, these models have relatively low predictive power (Aldrich and Simon, 1986:277–278). A major omission that may account for this poor performance is failure to take account of social environment and peer connections. For if these factors are important, their omission represents a misspecification of the voting process (Duncan and Stenbeck, 1988:33).

The omission of interpersonal relationships results in misspecification of a turnout model in two important ways. From a statistical perspective, the standard assumption of independent observations is violated. This is especially serious when a cluster sampling design includes all adults within a sampled household, such as the Current Population Surveys analyzed by Wolfinger and Rosenstone (1980). From a theoretical perspective, omitting important variables from a model usually results in biased parameter estimates for the included variables (Pindyck and Rubinfeld, 1981:128–130).

Choice of a model should reflect processes theorized to generate turnout, rather than the data and estimation procedures conveniently available. It is not unreasonable to theorize that associates' voting behaviors are interdependent, and that people are selective in picking their associates. Consequently, a nonrecursive model is appropriate, even though this would require better data and estimation procedures. Testing a psychologically oriented model, for example, would require collecting attitudinal as well as behavioral measures for respondents' spouses and other close associates. Simultaneous equation estimation procedures need to be developed for categorical dependent variables,[8] and initial efforts to model personal influence over time in dyads and networks need to be extended (Duncan, 1985:119–121; Friedkin and Johnsen, forthcoming).

In summary, the results support Glaser's (1959) contention that explanations of voting turnout cannot ignore interpersonal relationships.

References

Aldrich, John H., and Dennis M. Simon (1986)
"Turnout in American national elections." Pp. 271–301 in Samuel Long (ed.), Research in Micropolitics, Vol. 1. Greenwich, CT: JAI Press.
Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee (1954)
Voting. Chicago: University of Chicago Press.
Campbell, Angus, Gerald Gurin, and Warren E. Miller (1954)
The Voter Decides. Evanston: Row, Peterson and Company.

---

[8] Wilson (1988) has shown that the conventional wisdom on simultaneous equation models for categorical data (Fienberg, 1980:133–134) is incorrect and that such models for categorical data are meaningful. However, he does not provide estimation procedures.

Duncan, Otis Dudley (1985)
"New light on the 16-fold table." American Journal of Sociology 91:88–128.

Duncan, Otis Dudley, and Magnus Stenbeck (1988)
"Panels and cohorts: Design and model in the study of voting turnout." Pp. 1–35 in Clifford C. Clogg (ed.), Sociological Methodology, 1988. Washington: American Sociological Association.

Fienberg, Stephen E. (1980)
The Analysis of Cross-Classified Categorical Data. 2d ed. Cambridge: MIT.

Friedkin, Noah E., and Eugene C. Johnsen (forthcoming)
"Social influence and opinions." Journal of Mathematical Sociology.

Glaser, William A. (1959)
"The family and voting turnout." Public Opinion Quarterly 23:563–570.

Grofman, Bernard (1987)
"Models of voting." Pp. 31–61 in Samuel Long (ed.), Research in Micropolitics, Vol. 2. Greenwich: JAI Press.

Hauser, Robert M., and David L. Featherman (1977)
The Process of Stratification. New York: Academic Press.

Huckfeldt, Robert, and John Sprague (1987)
"Networks in context: The social flow of political information." American Political Science Review 81:1197–1216.

——— (1988)
"Choice, social structure, and political information: The informational coercion of minorities." American Journal of Political Science 32:467–482.

Jennings, M. Kent, and Gregory B. Markus (1988)
"Political involvement in the later years: A longitudinal survey." American Journal of Political Science 32:302–316.

Kinder, Donald R., and David O. Sears (1985)
"Public opinion and political action." Pp. 659–741 in Gardner Lindzey and Elliot Aronson (eds.), The Handbook of Social Psychology, 3d ed., Vol. 2. New York: Random House.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet (1948)
The People's Choice. 2d ed. New York: Columbia University Press.

Pindyck, Robert S., and Daniel L. Rubinfeld (1981)
Econometric Models and Economic Forecasts. 2d ed. New York: McGraw-Hill.

Silver, Brian D., Paul R. Abramson, and Barbara A. Anderson (1986)
"Who overreports voting?" American Political Science Review 80:613–624.

U.S. Bureau of the Census (1982)
"Voting and registration in the election of November 1980." Current Population Reports, ser. P-20, no. 370. Washington: U.S. Government Printing Office.

——— (1983)
"Voting and registration in the election of November 1982." Current Population Reports, ser. P-20, no. 383. Washington: U.S. Government Printing Office.

——— (1986)
"Voting and registration in the election of November 1984." Current Population Reports, ser. P-20, no. 405. Washington: U.S. Government Printing Office.

Wilson, Thomas P. (1988)
"Structural-equation models for categorical data: Non-recursive models and reduced forms." Santa Barbara: Department of Sociology, University of California.

Wolfinger, Raymond E., and Steven J. Rosenstone (1980)
Who Votes? New Haven: Yale University Press.

WHITE REACTIONS TO BLACK CANDIDATES

WHEN DOES RACE MATTER?

JACK CITRIN
DONALD PHILIP GREEN AND
DAVID O. SEARS

Abstract This paper examines the role of race in elections where one of the candidates is black. Using the 1982 California gubernatorial election between Tom Bradley and George Deukmejian as a case study, the paper shows that whereas racial attitudes were a significant influence on the voting decisions of whites, Bradley's background did not stimulate an unusual level of racially motivated behavior. The paper argues that the impact of a candidate's race on voting depends on a number of contextual factors, including his prior record and campaign style. The paper also proposes a technique for comparing the results of biracial elections with contests where all the candidates are white as a method for estimating the level of racial voting.

After several decades of apparent growth in racial tolerance, how do whites react when a black runs for a major office? The recent candidacies of Jesse Jackson, Tom Bradley, Wilson Goode, and Harold Washington for prominent national, state, and local positions has renewed interest in the extent to which a politician's race affects citizens' voting decisions. This paper analyzes the California gubernatorial election of 1982 in which Tom Bradley, the black Democrat, lost to George

JACK CITRIN is Associate Professor of Political Science at the University of California at Berkeley. DONALD P. GREEN is Assistant Professor of Political Science at Yale University. DAVID O. SEARS is Professor of Psychology and Political Science at UCLA. The authors are grateful to Samm Mcmurdo, Beth Reingold, and Evelyn Walters for their research assistance and to Herbert McClosky and Robert Luskin for many helpful criticisms and suggestions. This research was supported by grants for the Institute of Governmental Studies and the Survey Research Center, University of California at Berkeley. The data was provided by the State Data Program of the Survey Research Center and full descriptions of the relevant surveys and precise wording of the questions analyzed are available from the program's data library.

Public Opinion Quarterly Volume 54:74–96 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

Deukmejian, the white, conservative Republican, by the slim margin of 100,000 votes out of the almost seven and a half million cast.

Race and Mass Politics

Recent research on the role of race in American mass politics has centered on three main theoretical issues. One question raised is the extent to which the voting decisions of whites are influenced by racial factors. Numerous studies have shown that race has a pervasive influence, affecting not only the reactions of whites to black candidates ranging from Edward Brooke (Becker and Heaton, 1967) and Bradley (Kinder and Sears, 1981) to Harold Washington (Baker and Kleppner, 1986) and Jesse Jackson (Sears, Citrin, and Kosterman, 1987) and their preferences on manifestly racial policies such as busing (Sears and Allen, 1984) and affirmative action (Sears, Huddy, and Shaffer, 1986), but also their opinions on ostensibly nonracial policies such as crime (Sears et al., 1980) or property tax cuts (Sears and Citrin, 1985).

A second issue revolves around the motivational basis of these racially driven effects. The main controversy here has centered on whether self-interest or realistic group conflicts are implicated, as opposed to racial prejudice operating in a purely affective way. The current consensus appears to be that *personal* self-interest is rarely an important factor in antiblack voting or policy preferences (Citrin and Green, 1990; Sears and Funk, 1988), with the jury still out on the role of group interest (Bobo, 1983; Sears, 1988; Schuman and Bobo, 1988).

The third issue concerns the nature of the antiblack feelings at work. Kinder and Sears (1981) and McConahay and Hough (1976) argued that a new form of racism, labeled "symbolic" or "modern," had supplanted "old-fashioned" or "redneck" stereotyping of blacks as inherently inferior as the most potent racial influence on whites' political preferences. There does appear to be some scholarly agreement that the common manifestations of racial prejudice have changed, but the scope and meaning of this change remain a matter of dispute (Sears, 1988; Kinder, 1986; Schuman, Steeh, and Bobo, 1985; Sniderman and Tetlock, 1986).

Our concern in the paper is limited to the question of the role of a candidate's race in triggering negative responses from white voters. One reason that the 1982 California gubernatorial contest is a particularly useful case study is the importance of the office at stake. Had Bradley won, he would have attained the highest political position of any black in American history. Another reason is that preelection polls consistently predicted that Bradley would win, leading some observers to take the actual outcome as evidence of substantial convert racism among white Californians (Henry, 1983). Bradley's prior record and campaign strategy provide another reason. By 1982 the mayor of Los Angeles had established a solid reputation as a moderate on racial matters, and his campaign deliberately avoided overt appeals to minority groups as their champion (Pettigrew and Alston, 1988).

The trajectory of Bradley's career thus provides an opportunity to assess how white voters adapted to a black candidate for governor who presented himself in a race-blind manner. If even so unthreatening a politician as Bradley were to provide widespread resistance, this would be strong evidence for the pervasive power of racial antagonisms in American elections. The hypothesis of this paper is that the personal attributes of the black candidate and the electoral context do condition the reactions of white voters. We argue that in Bradley's case this resulted in the diminished importance of his race relative to other political considerations as the basis of opposition to him.

Cues for Racial Voting

If racism means the differential treatment of individuals simply on the basis of their racial group membership (Katz, 1976), then one meaning of racial voting is the rejection of black candidates just because of their race.[1] A second meaning focuses on the beliefs rather than the race of a candidate and defines racial voting as a preference for any candidate perceived to oppose the distinctive political goals of blacks.

As Wolfinger (1974:49) has argued, for group-related attitudes to influence voting behavior they must be engaged by situational stimuli. The nomination of a black is one cue of the ethnic relevance of voting decisions, especially when candidates make racial solidarity a prominent campaign theme. More generally, previous researchers have argued (Key, 1949; Wolfinger, 1974; Pettigrew, 1976) that the perception of racial threat is critical in provoking negative reactions to black candidates among whites. Among the factors that seem to enhance such sentiments are the size of the black population, the history of race relations in the community, and the salience of racial issues in the campaign. In addition, whether the election is partisan or not should matter. A Democratic party identification may check the inclination of racially intolerant whites to oppose a black candidate.

---

[1] This paper is concerned with the role of Bradley's race in determining whether or not whites voted for him. A comprehensive analysis would deal with changes in turnout as well as electoral choice among *both* racial groups. But while we briefly address the question of black mobilization below, the paucity of black respondents in the surveys we analyze limits our study to the behavior of whites.

One possible stimulus for anti-Bradley voting among whites in 1982 was that this marked the first genuine opportunity for a black to become governor of a major state. In mayoral elections at least, the first serious candidacy of a black increases concern among whites regarding the allocation of power between the races (Pettigrew, 1976). Another reason for expecting Bradley's race to count against him is that the political mood in California at the time was largely unsympathetic to the political demands of minority groups. The feeling that government was already doing too much for blacks contributed to support for the tax revolt (Sears and Citrin, 1985), for example, and a constitutional amendment that forbade forced busing of schoolchildren had won overwhelming approval at the polls in 1979.

Other features of the gubernatorial contest in 1982, however, seemed likely to mitigate the impact of race. Although racial fears had influenced voting in the Los Angeles mayoral election of 1969, in which Bradley first ran against Sam Yorty (Kinder and Sears, 1981; Hahn, Klingemann, and Pachon, 1976), by 1982 he was an uncontroversial figure whose entire political career bespoke caution and accommodation. In Los Angeles he had the solid support of the white establishment, and there was no serious opposition to him in the Democratic gubernatorial primary.

Aside from his reassuring personal qualities, Bradley campaigned as a moderate, stressing his fiscal conservatism and background in law enforcement in the hope of decreasing the defection of white Democrats. Deukmejian also avoided explicit discussion of racial policies, although some observers have characterized his "tough" stand on law and order issues as subtly racist, presumably because a disproportionate number of those receiving the stiffer punishments he advocated would be black.

Bradley's race did surface as a campaign issue in early October when Bill Roberts, Deukmejian's campaign manager, told reporters that covert antiblack sentiment would swing a close election to his candidate. A few days later Deukmejian accepted Roberts's resignation, ostensibly because the statement itself might be interpreted as a beacon for racist voters. His new campaign managers advertised Deukmejian in mailings and television spots as a governor for "all" Californians, a message that also may have racist undertones (Pettigrew and Alston, 1988).

Clearly, racial themes were not entirely absent from the Bradley–Deukmejian contest. But their muted quality stands in striking contrast to the mayoral elections in Chicago in 1983 and 1987, which were explicitly framed as a conflict between blacks and whites. So the question we return to is this: How does race influence voting behavior when neither candidate, black or white, makes the issue salient?

Measuring Racial Voting

The simple comparison of how blacks and whites voted does not identify the motivational basis of electoral choice and thus is an inadequate index of the role of racism. A more acceptable method for determining the impact of racial attitudes on voting is a multivariate analysis in which one controls for the influence of alternative causes such as party loyalty or opinions or nonracial issues (Sears, Hensler, and Speer, 1979; Sears, Citrin, and Kosterman, 1987). Thus, our assessment of racial voting among whites is based on the statistical effect of their racial attitudes on their electoral decisions.

As noted above, there is a crucial distinction between the impact of racial attitudes on voting and the influence of the candidate's race per se. Indeed, given that the Democratic party, nationally and in California, is widely perceived as more liberal on civil rights issues and more likely to support the political aspirations of blacks and other minority groups (Carmines and Stimson, 1980), substantial racial voting is to be expected even when both candidates are white. To establish whether Bradley's race caused whites to vote against (or for) him therefore requires some standard of comparison for the "effect" coefficient described above.

The ideal comparison would be to contrast the role of racial attitudes in the Bradley–Deukmejian contest with their influence in an election where the Democratic candidate was a white clone of Bradley. But since this ideal is inherently unattainable, we worked with the best available alternative—the other statewide elections in 1982.[2] Thus, we consider whether racial voting varies systematically across these elections and infer the impact of Bradley's race from the *relative* importance of antiblack opinions in the gubernatorial election. Accordingly, the statistical test is whether racial attitudes had a larger effect on voting for governor than on electoral choices in the contests without a black candidate.

Table 1 summarizes the main features of the four elections we analyze, namely the gubernatorial contest and the elections for U.S. senator, lieutenant governor, and state superintendent of schools. As this table indicates, the differences between the Bradley–Deukmejian election and the others vary from one paired comparison to the next.

---

[2] Another possible baseline is the historical record: we might compute an expected Democratic vote based on the outcomes of recent elections in which there were no black candidates, measure the deviation from this "normal vote," and attribute this difference to the short-run impact of Bradley's race (Campbell et al., 1964; Boyd, 1972). We rejected this technique here, partly because in a period characterized by electoral volatility it is risky to assume, even for heuristic purposes, that there is an underlying "normal vote" and partly because there is no method for disentangling the unique contribution of any of several short-run factors.

Table 1. Characteristics of the 1982 California Election

	Race of Candidates	Partisan Contest	Salience of Race Issues	Level of Interest
Governor (Bradley vs. Deukmejian)	Biracial	Yes	Low	High
U.S. Senate (Brown vs. Wilson)	Both white	Yes	Low	High
Lt. governor (McCarthy vs. Hallett)	Both white	Yes	Low	Low
Sup't. schools (Riles vs. Honig)	Biracial	No	Low	Low

For example, the senatorial election pitted Jerry Brown, the flamboyant white Democrat, against Pete Wilson, a moderately conservative Republican who ultimately won. During Brown's tenure as governor of the state, the appointment of blacks, Hispanics, and women was a prominent policy. Thus although his campaign, like Bradley's, did not emphasize racial issues, his prior record led voters to regard him as favorable to black political needs and demands. In comparing the underpinnings of support for Bradley and Brown, therefore, we "control" for the candidates' party, ideology, and nonincumbent status but not for their personal popularity or the level of government of the office being contested.

In contrast to the senatorial and gubernatorial elections, the contests for lieutenant governor and state superintendent of schools were both low-key events. Neither of the incumbent candidates, Leo McCarthy and Wilson Riles, carried the negative personal baggage of Jerry Brown. McCarthy, the white Democrat, resembled Bradley in his conventional liberalism and restrained personality. But unlike Bradley, he won. However, Riles, a black, lost the nominally nonpartisan contest for state superintendent of schools to a white opponent. In this campaign, the performance of California's public schoolchildren on standardized test scores, not manifestly a racial matter, was the dominant issue.

The fact that the four elections we compare varied on theoretically relevant characteristics enables us to specify more precisely the conditions under which a candidate's race matters. Specifically, we can go beyond the simple comparison of biracial and all-white contests to consider the relevance of contextual factors such as the prominence of the office being contested, whether the election is partisan in nature, and the personal attributes of the black candidate.

Data

This study is based on a secondary analysis of two preelection surveys of registered voters.[3] The polls we analyze are a *Los Angeles Times* survey conducted 10–14 October 1982, two and a half weeks before the 2 November election, and a California Poll conducted by the Field Institute on 24–31 October.[4] It is possible, therefore, that there were unmeasured changes in attitudes on racial and other issues in the final days of the campaign. Indeed, both samples included a substantial number of respondents who said they were still undecided about how

---

3. Pettigrew and Alston (1988) analyzed the role of racism in these elections by relying mainly on countywide voting statistics. This approach has the advantage of treating votes rather than voting intentions as the dependent variable and thus overcomes some of the biases inherent in self-reported information. On the other hand, it confronts the formidable difficulty of making valid inferences about the motivations of individuals from aggregate data. This problem is acute in the present case, since California's counties are so varied in size and social composition. Both aggregate and survey data would seem to be valuable, therefore, each with its own strength and fallibility.

4. Both surveys employed telephone interviewing and sampled a representative cross-section of the California population over eighteen. Both also used random digit dialing techniques in which numbers are generated in proportion to local prefix allocation in order to remove the biases due to the absence of unlisted numbers from the directories used to select the initial clusters. In addition, the California Poll employed a multistage weighting procedure to ensure the representativeness of the sample with respect to age, sex, and party registration. The *Los Angeles Times* survey employed a screening question to confine interviewing to respondents self-reported as registered to vote, while the California Poll fully interviewed only registered voters who said they were "absolutely certain" to vote in the upcoming election.

The *Los Angeles Times* poll made three callbacks in an effort to reach a respondent; their sample of 1,578 adult Californians yielded 833 registered voters whose answers are analyzed here. Interviewers for the California Poll made four callbacks. Neither organization attempted to convert respondents who refused to be interviewed. Unfortunately, neither organization was able to provide the exact number of nonresponses of different types for the particular surveys analyzed here.

Pettigrew and Alston (1988:12) list a number of other surveys of opinion on the Bradley–Deukmejian contest. These preference polls were conducted by either Teichner Associates or Tarrance & Associates. However, to the best of our knowledge, they are not generally available for reanalysis and do not include measures of racial attitudes. Nevertheless, the distributions of voting intentions reported by these polls are consistent with our own results. They show that Bradley began the campaign with a solid advantage that substantially eroded by election day. We should also note that the exit poll conducted by the Field Institute has not been released for secondary analysis. The surveys we analyze here, therefore, are the only generally available individual-level data that can relate opinions on racial issues to voting intentions in the 1982 statewide elections.

to vote. Clearly, it would have been preferable to be able to analyze survey data collected in the last few days of the campaign.

Each survey contained several questions that tap racial attitudes. Table 2 lists these questions and provides the marginal distributions of the responses. The questions contained in the *Los Angeles Times* survey (*LAT*) can be subdivided into those that tap resentment of governmental favoritism toward blacks and others that attribute stereotypically unfavorable personal characteristics to blacks as a group.[5] The core of our analysis employs these indicators of racial attitudes as predictors of voting behavior in the four statewide elections.

Results

How did white voters in 1982 evaluate Tom Bradley? By historical standards, Bradley's candidacy did not induce an abnormally high rate of partisan defection. According to the CBS News exit poll, he obtained 44.9% of the votes of all whites and 79% of the votes of those registered as Democrats (Baker and Kleppner, 1986:227). California Poll data show that Jerry Brown won 81% of the votes of white Democrats in each of his two victorious gubernatorial campaigns and that Pat Brown retained the support of only 71% of this group when he was defeated by Ronald Reagan in 1966.

Bradley had a highly favorable personal image. The *LAT* survey asked respondents to rate the candidates for major statewide offices on a scale of 1 (most unfavorable) to 4 (most favorable). *Among whites*, Bradley's mean rating was 2.99, with 79% rating him either favorably or very favorably. This was a more positive evaluation than either Deukmejian (2.78) or the two white senatorial candidates. Wilson received a 2.69 mean rating and Brown a much lower 2.12, with only 38% of the white electorate evaluating him favorably. The California Poll conducted just before the election found the identical pattern of ratings using a scale from 1 (low) to 10 (high) to rate the candidates. In this survey, among whites who were registered and said they intended to

[5] The former category resembles the items described elsewhere as measures of "symbolic" or "modern" racism, an attitude that fuses antiblack affect and conservative morality (Sears, 1987; McConahay, 1986), but it may also tap conceptions of group interest or general feelings about the role of government (Schuman and Bobo, 1988). The second category has not been described in those terms in the past. Indeed, unflattering stereotypes of blacks are usually treated as indicators of "old-fashioned racism" (McConahay, 1986). However, it is plain that the issues of crime and violence often stand as symbols for racial matters. For this reason, it would be hard to make a case that differences in the effect of these measures bear on the theoretical disputes described above concerning the exact nature of political racism. To repeat, in this paper we are concerned with estimating the effects of racial attitudes in a global sense.

Table 2. Racial Opinions Among White Californian Voters, October 1982

Government Treatment of Blacks

The government should not make any special effort to help Blacks and other racial minorities because they should help themselves. (CP)

Agree strongly	20.4%
Agree somewhat	13.8
Disagree somewhat	24.1
Disagree strongly	35.5
Don't know	6.2
	100.0
N	730

Over the past few years, Blacks have been getting more favorable treatment from government than other races have. (CP)

Agree strongly	20.5%
Agree somewhat	21.8
Disagree somewhat	24.8
Disagree strongly	21.9
Don't know	11.0
	100.0
N	730

Do you think that government has paid too much attention to blacks and other minority groups . . . or about the right amount . . . or do you think the government hasn't paid enough attention to blacks and other minority groups? (*LAT*)

Too much attention	32.8%
Right amount	44.4
Not enough attention	15.4
Don't know	6.3
Refused	1.1
	100.0
N	833

Table 2. (Continued)

How do you think the government in Sacramento should treat blacks and other minorities? Do you think it should make every effort to improve their economic condition, or do you think it should make no special effort to help them? (*LAT*)

Make every effort	46.0%
Make no special effort	43.3
Don't know	8.8
Refused	1.9
	100.0
N	833

Perceived Traits of Blacks

Which of the following groups are the least law-abiding: Blacks, Italians, Japanese-Americans, Jews, or Mexican Americans—or do you think they are all pretty much equal in that respect? (*LAT*)

All equal	34.3%
Blacks	24.2
Hispanics	20.3
Other	1.8
Not sure	15.2
Refused	4.2
	100.0
N	833

Which of the following groups are more prone to violence: Blacks, Italians, Japanese-Americans, Jews, or Mexican Americans—or do you think they are all pretty much equal in that respect? (*LAT*)

All equal	31.8%
Blacks	23.3
Hispanics	25.5
Other	0.8
Not sure	13.2
Refused	5.4
	100.0
N	833

Sources: CP = California Poll with percentages based on the total sample *N* of 730. *LAT* = *Los Angeles Times* poll with percentages based on the total sample *N* of 833.

vote, Deukmejian and Bradley had almost identical ratings (6.3 vs. 6.1). But once again, Bradley was much more favorably viewed than Jerry Brown (4.4).

Turning now to public opinion on racial matters, only 3% of the respondents in a September 1982 California Poll mentioned Bradley's ethnicity as a reason they might vote against him (Henry, 1983). Moreover, a majority of these respondents were Republicans who were unlikely to vote for a Democratic candidate of either race. However, people might be reluctant to state openly that they would not vote for someone just because he is black, so this kind of question may understate the level of negative racial reactions.

Indeed, the data in Table 2 show that unfavorable images of blacks and opposition to their political aspirations were much more widespread than the positive statements about Bradley himself might imply. For example, 23% of whites in the *LAT* survey felt that blacks were more prone to violence than any other racial or ethnic group, 43% believed that the state government should make no special effort to help blacks and other racial minorities, and 33% complained that blacks had been receiving too much attention from government.

Tables 3 and 4 present the relationships between racial feelings and voting intentions in the 1982 statewide contests, controlling for the respondent's party registration.[6] These fundamental points emerge:

1. Negative feelings about blacks consistently increased opposition to Bradley. For example, only 32% of the whites in the *LAT* survey who felt government had been paying too much attention to blacks and other minorities said they would vote for Bradley compared to 79% of those who believed government was not doing enough for these groups. In the CP survey, only 39% of the whites who agreed strongly that "the government should not make any special effort to help blacks" said they would vote for Bradley compared to 61% of those who disagreed strongly with this statement.
2. The relationships between racial attitudes and voting intentions in the elections where both candidates were white and in the biracial Bradley–Deukmejian contest were very similar. If anything, the data suggest that Bradley was more popular than Brown even among those expressing antiblack sentiments. For example, if one compares whites in the CP survey who agreed strongly that the government should not make any special effort to help blacks with those who disagreed strongly, the *differential* in support for Brown among white Democrats was 26

---

[6] The same pattern of findings emerges when racial beliefs are related to the respondents' ratings of the candidates. We do not report these data here for the sake of brevity.

Table 3. Racial Beliefs and Electoral Choice Among Decided Voters, by Party, *Los Angeles Times* Poll

Is government paying too much attention to blacks?

	Democrats			Republicans		
	Too Much	Right Amount	Not Enough	Too Much	Right Amount	Not Enough
Vote for Bradley	61%	81%	93%	11%	28%	33%
	(74)[a]	(134)	(72)	(97)	(126)	(24)
Vote for Brown	55%	62%	93%	8%	10%	26%
	(69)	(133)	(67)	(109)	(126)	(23)

Should Sacramento make every effort to help blacks?

	Democrats		Republicans	
	No Special Effort	Make Every Effort	No Special Effort	Make Every Effort
Vote for Bradley	66%	86%	19%	25%
	(121)	(154)	(120)	(118)
Vote for Brown	56%	77%	8%	14%
	(115)	(149)	(126)	(118)

Which group is less law-abiding?

	Democrats			Republicans		
	Blacks	Hispanics	All Equal	Blacks	Hispanics	All Equal
Vote for Bradley	73%	78%	76%	20%	25%	37%
	(71)	(67)	(105)	(81)	(59)	(95)
Vote for Brown	67%	63%	68%	12%	10%	13%
	(67)	(62)	(108)	(83)	(62)	(98)

Which group is more prone to violence?

	Democrats			Republicans		
	Blacks	Hispanics	All Equal	Blacks	Hispanics	All Equal
Vote for Bradley	73%	73%	82%	18%	27%	20%
	(70)	(79)	(98)	(66)	(77)	(98)
Vote for Brown	70%	69%	63%	8%	16%	11%
	(71)	(74)	(97)	(73)	(74)	(101)

SOURCE: *Los Angeles Times* Poll, October 1982.

[a] The numbers of cases in parentheses refer to the base for percentaging. For example, 74 refers to the number of Democrats in the *LAT* poll who said that government is paying too much attention to blacks; 61% of this group voted for Bradley.

Table 4. Racial Beliefs and Electoral Choice Among Decided Voters, by Party, California Poll

	Democrats				Republicans			
	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree
Government should *not* make any special effort to help blacks.								
Vote for Bradley	62%	63%	69%	78%	18%	15%	19%	27%
	(47)	(27)	(71)	(128)	(66)	(47)	(70)	(77)
Vote for Brown	50%	67%	66%	76%	6%	9%	8%	26%
	(46)	(30)	(67)	(127)	(65)	(45)	(73)	(84)
Vote for McCarthy	36%	37%	69%	67%	16%	36%	32%	43%
	(45)	(27)	(61)	(117)	(57)	(36)	(66)	(74)
Vote for Riles	67%	79%	79%	76%	15%	27%	12%	29%
	(46)	(24)	(57)	(117)	(59)	(41)	(65)	(55)
Blacks have been getting more favorable treatment from government.								
Vote for Bradley	55%	64%	73%	18%	22%	23%	15%	32%
	(47)	(55)	(64)	(79)	(72)	(62)	(71)	(38)
Vote for Brown	47%	58%	69%	82%	10%	13%	8%	32%
	(47)	(53)	(64)	(82)	(72)	(64)	(71)	(41)
Vote for McCarthy	42%	62%	53%	67%	28%	31%	36%	40%
	(45)	(53)	(57)	(70)	(65)	(52)	(64)	(35)
Vote for Riles	52%	77%	87%	77%	18%	23%	16%	35%
	(44)	(44)	(62)	(74)	(62)	(53)	(67)	(26)

SOURCE: California Poll, October 1982.

percentage points. The comparable differences in the level of support for Bradley and Leo McCarthy were 16 and 10 percentage points respectively.

3. The impact of racial beliefs on voting intentions was more potent in the biracial Riles–Honig contest than in the gubernatorial contest. This, we believe, reflects the stronger influence of party affiliation in the latter election and the fact the electorate was much more informed about Bradley's record and policy positions than Riles's.
4. The *LAT* data indicate that the negative stereotypes about the personal qualities of blacks were much more weakly related to voting intentions than resentment about the government's perceived favoritism toward racial monorities. Among white Democrats, for example, those who felt the government was paying too much attention to blacks were less likely to vote for Bradley than those who felt it was not paying enough attention by a margin of 36 percentage points. But the differential in support for Bradley among those who believed blacks were more prone to violence than any other ethnic group and those who felt all groups were equal in this regard was only 9 percentage points.

What, then, was the impact of race in the Bradley–Deukmejian contest in 1982? The evidence suggests that the answer depends on which facet of racial voting one examines. It is clear that overt resentment about governmental support for black demands reduced support for Bradley. However, racial attitudes had a similar effect on voting for Brown and McCarthy, which implies that Bradley's race did not, *in and of itself*, cause many whites to vote against him.

To estimate the role of racial attitudes more precisely, Table 5 summarizes the results of multivariate analyses that include preferences on other issues as predictors of how people voted. Because voting intention, the dependent variable, is dichotomous, the estimation procedure is probit rather than ordinary least squares.[7] Table 5 reports the results for all four statewide elections and employs data from both the *LAT* and CP surveys. The model includes the respondent's age, sex, level of education, party affiliation, racial beliefs, and opinions on several policy questions that have no *manifest* linkage to race as predictors of voting decisions. To facilitate comparison of their effects, all the pre-

---

[7] For the gubernatorial and senatorial elections, the analysis was replicated with a net candidate evaluation measure derived by subtracting the respondent's rating of Deukmejian (Wilson) from his rating of Bradley (Brown) as the dependent variable. The estimation procedure employed in this instance was OLS regression. In every case, the results were fundamentally the same as those reported in Table 5.

Table 5. Probit Estimates of Voting Intentions in Four Elections, for Two Models of the Vote

California Poll

	Governor		U.S. Senator		Superintendent		Lt. Governor	
	Probit	*r*	Probit	*r*	Probit	*r*	Probit	*r*
Party	1.19*	.48	1.36*	.52	.45*	.24	1.30*	.52
Education	.11*	.13	.11	.10	.10	.11	.03	.03
Gender	.15	.00	.18	.02	.05	.03	.10	.10
Age	−.05	−.07	−.03	.05	−.01	−.01	.01	.01
Black treatment	.31	.23	.72*	.29	.56*	.23	.19	.21
Welfare	.04	.19	.24	.17	.15	.13	.09	.16
Opinion on gun control (Prop. 15)	.36*	.26	.33*	.25	.23	.20	.28*	.22
Gun ownership	.10	.16	.10	.12	.05	.09	.10	.15
Size of gov't	.71*	.33	1.21*	.42	.68*	.27	.79*	.35
N	500		505		448		437	

Los Angeles Times Poll

	Governor		U.S. Senator	
	Probit	*r*	Probit	*r*
Party	1.24*	.55	1.14*	.56
Education	.02	.04	.07	.04
Gender	.28	.10	−.06	−.06
Age	−.16*	−.18	.01	.07
Black treatment	1.05*	.33	.98*	.34
Back character	−.02	.05	−.01	.03
Opinion on gun control (Prop. 15)	.33*	.42	.28	.23
Political ideology	1.23*	.15	1.94*	.47
Salience of crime	.93	.15	.14	.03
N	440		433	

NOTES: Entries are probit coefficients and Pearson correlations. Dependent variables are scored with the Democratic candidate (and Wilson Riles) as 1 and the Republican (and Bill Honig) as 0. Party is scored 0 if the respondent is a Republican, 1 if the respondent is a Democrat. Education is scored 1 for those with less than a high school diploma, 2 for those with a high school degree, 3 for those with some college, 4 for those with a college degree, and 5 for those with some graduate training. Gender is scored 1 if the respondent is female. Age is divided into four categories: ages 18–35 = 1, 36–50 = 2, 51–65 = 3, 66 and up = 4. For the California Poll data, the welfare item is divided into four categories, from strongly agree (0) to strongly disagree (1). Gun control is scored 0 if the respondent favored Proposition 15, 1 if opposed, .5 if undecided. Gun ownership is a dummy variable scored zero if either guns or rifles were present in the respondent's household. The size of government index contains four categories and is scored in the "big government" (or pro-Democratic) direction. For the *Los Angeles Times* data, self-designated political ideology has five categories, ranging from strong conservative (0) to strong liberal (1). Crime salience represents the respondent's assessment of the importance of crime as a political issue, ranging from 0 to 1 in order of decreasing importance. The black treatment and black character indices, described in detail in the text, range from 0 to 1.

* $p < .05$.

dictors excepting the demographic variables were rescored to vary between 0 and 1.

For respondents in both surveys, racial attitudes were measured by a black treatment index constructed by combining answers to the questions about the amount of attention black political demands should receive and about whether blacks deserved special treatment from state government. For respondents in the *LAT* survey, we also constructed a black character index by adding responses to the questions about the perceived violence and law-abidingness of diverse ethnic groups.[8] These two indices of racial feelings correlated at .18.

Because the *LAT* and CP surveys did not include the same questions, the predictive equations for the two sets of data are not identical. In the *LAT* survey, the measures of ostensibly nonracial issues were a question about the importance of the crime issue to the respondent, the respondent's ideological self-identification, and the respondent's intended vote on handgun control (Proposition 15). When using the CP survey, we included as measures of nonracial issue positions the respondent's intended vote on handgun control, a two-item index assessing beliefs about the proper size of government, and responses to a question about whether people receiving welfare truly needed it.

Previous research by Sears and Citrin (1985) showed racial attitudes are related to opinions on taxing and spending issues. In the present study, preferences on the nonracial issues included in the predictive model consistently were associated with opinions about racial minorities. For example, in the CP survey the black treatment index correlated .21 with how one intended to vote on the handgun control initiative, .23 with scores on the index measuring attitudes toward the proper size of government, and .36 with the question about the need for government spending on welfare.

Under these circumstances, it is not surprising that the "race effect" reported in Tables 3 and 4 is substantially reduced in the multivariate model presented in Table 5. The substantive meaning of this result, however, is ambiguous. If one regards conservative preferences on issues such as gun control or government spending as subtly, subliminally, or symbolically racist, then some portion of their observed influence on the vote should be treated as a "race effect."

In the *LAT* survey, explicitly racial attitudes retained a statistically significant influence on voting for Bradley even after we controlled for party registration, ideological self-identification, and opinions on crime and the sale of handguns. In the CP data, however, the black treatment variable failed to have a significant independent effect on voting for Bradley in the multivariate model. The predictors with statistically significant probit coefficients in the CP data were opinions about the proper size of government, attitudes toward gun control, and ideological self-identification, all part of the syndrome of beliefs and symbolic attachments that incorporates attitudes toward blacks. Thus, while we are not in this paper attempting to derive a precise estimate of the total effect of racial attitudes on the gubernatorial vote, we believe this is greater than the minimalist figure provided by the black treatment coefficient alone.[9]

---

[8] For each item, respondents who named blacks or Mexican Americans as more prone to violence (or less law-abiding) were given a score of 0, those who named some other group were scored .5, and those who said all groups were equal or refused to answer were coded as 1.

As we have repeatedly emphasized, the critical evidence for the question of whether Bradley's race, rather than his party affiliation or perceived liberalism, cost him white votes is the *relative* explanatory power of antiblack sentiments in predicting voting decisions across the four statewide elections. Table 5 again shows that the overall effect of racial attitudes, however they are computed, is no stronger in the biracial Bradley–Deukmejian contest than in the contests without a black candidate. In the *LAT* survey, the black treatment coefficient is almost the same in the equations predicting gubernatorial and senatorial votes. In the CP data, this measure is a significant predictor of voting for Brown, but not for either McCarthy or, as noted above, Bradley. More generally, the structure of support for the Democratic candidates in the three partisan contests is very similar, suggesting that the differences in their outcomes may well be explained by the relative quality of the opposition's candidates.

The results for the contest between Riles and Honig for the office of state superintendent of schools provide an important insight into the way the electoral context shapes the influence of race on voting. The candidates in this election were similar in several respects. Both were professional educators. Neither was publicly identified with a particular political party or ideological camp. Indeed, neither was widely known among the electorate. The main differences between them were that Riles was the incumbent and that he was black.

Given that the election for superintendent of schools is nominally a nonpartisan affair, it is understandable that the estimated coefficient for party affiliation is much smaller in this instance.[10] What is striking is the significant influence of antiblack sentiment in this biracial contest. Like Bradley, Riles projected a moderate and conservative personal image. His prior record could hardly be viewed as inimical to "white interests," and his campaign ignored racial issues. Yet the belief that government was doing too much to help blacks gave a boost to his white challenger. The implication is that in the absence of a partisan anchor and little detailed information about the candidates, race becomes a more salient cue for voting decisions.

---

[9] Not even the strongest statement of the symbolic racism perspective considers conservatism on social issues to be the equivalent of racial prejudice. As Schuman and Bobo (1988) point out, preferences on these issues have varied attitudinal sources, of which antiblack prejudice is one.

[10] For all four elections, we also examined the possibility that racial attitudes were a more important determinant of electoral choice among Democrats than Republicans by modifying the six probit equations presented in Table 5 to include a multiplicative term to

Discussion

Throughout this paper we have distinguished between two meanings of racial voting. The first refers to the impact of racial attitudes on electoral decisions. On this point, our evidence points to the pervasive impact of race in American politics. Resentment over the extent of the government's attentiveness to the demands of blacks had a consistent influence on voting decisions in California in 1982. Racial attitudes affected voting largely through their association with partisanship and with opinions on other so-called social or cultural issues. In California, as elsewhere in the nation, antiblack feelings pushed one toward Republican candidates, even if one were registered as a Democrat. Negative feelings toward minorities were one element in a cluster of conservative beliefs—opposition to handgun control, to "big" government, to "soft" judges, and to progressive "social engineering." Directly or indirectly then, underlying racial attitudes contributed to the defeat of Tom Bradley, Jerry Brown, and Wilson Riles.

The second meaning of racial voting refers to the specific role of a candidate's race in engaging group identifications. From this perspective, our analysis indicated that Bradley's race was not an important catalyst for racial voting in 1982. This conclusion diverges from previous research on this election. Specifically, Pettigrew and Alston (1988) conceptualize and measure racial voting in a somewhat different manner. They report a significant relationship at the county level between voting for a 1972 antibusing ballot initiative and greater support for McCarthy than Bradley and infer from this that many whites voted against the Democratic gubernatorial candidate simply because he was black.

Doubtless there is a segment of the California electorate that would

---

measure the interaction between party affiliation and racial sentiments. In no case did the size of the coefficient for this interaction term exceed the standard error. In this sense, the results of the multivariate model replicate the findings reported in the three-way cross-tabulation of Tables 3 and 4.

never vote for a black. It is extremely unlikely, however, that many in this group would have voted for any white liberal Democrat in 1982. Bradley's soothing approach, his reluctance to portray himself as a spokesman for the interests of blacks, and his emphasis of his managerial abilities apparently succeeded in keeping the defection of white Democrats to ordinary levels.[11] In this sense, his race did not greatly hurt him, although in a very close election the loss of even a small number of votes for this reason might have made the difference between victory and defeat.

But one might also ask why a popular black with such a moderate record and comforting image suffered even this "normal" loss of white votes. One admittedly speculative answer is that whatever their objective record, black candidates are stereotyped as liberals who are favorable to "big" government, "soft" on crime, and so forth. Thus, to the unknown extent that white voters attributed certain unpopular values to Bradley on the basis of his ethnicity alone, a form of racism contributed to his defeat.

The perceived need of a black candidate to assuage the anxieties of white voters may have hurt Bradley in another way too. Whether because he downplayed the issue of his race or not, Bradley failed to mobilize blacks. A February 1983 report by the Field Institute used surveys to estimate turnout and found that the level of voting among blacks was average in 1982. And the official records of the secretary of state in California show that in counties where blacks comprised more than 10% of the population turnout for the gubernatorial contest declined from 68.2% in 1978 to 66.7% in 1982. Given the closeness of the election in 1982, a modest increase in turnout among blacks would, other things being equal, have changed the outcome.

COVERT RACISM?

We have not yet considered the possibility that the social desirability of expressing tolerant opinions led white respondents to conceal their negative feelings about blacks. If covert racism were pervasive, distorting our measurement of racial attitudes or voting intention, we may have underestimated the impact of race in the Bradley–Deukmejian election.

While the existence of even a relatively small number of people who deliberately misrepresented their preferences may have served to falsify the pollsters' preelection predictions concerning the Bradley–Deukmejian contest, we do not believe that covert racism substantially attenuated the parameter estimates presented in Table 5. First, the polls we analyzed were conducted by telephone and did not concentrate on racial issues, thereby reducing the chances of eliciting the social desirability response set. Second, respondents did not appear reluctant to express antiblack sentiments. When one-third of the white respondents freely expressed the view that government was doing too much for racial minorities, it seems unlikely that others felt intense pressure to conform to a social norm of racial liberalism. Finally and most importantly, we have shown that the candidate's race did seem to matter in the contest between Riles and Honig. It is difficult to see why covert racism should have shaped answers to questions about Bradley but not about Riles.[12]

---

[11] Additional support for this conclusion comes from a comparison of the present results with the study of Bradley's mayoral elections conducted by Kinder and Sears (1981). That analysis indicated that the influence of symbolic racism on opposition to Bradley diminished between 1969 and 1973. In the later election racial issues were less prominent in the campaign and Bradley was much better known.

CONTEXT AND CANDIDATES

Throughout this paper we have contended that the influence of race on voting varies with the electoral context, broadly defined to include the black candidate's personal attributes and campaign strategy. In a state where less than 10% of the voters were black, Tom Bradley made a deliberate effort to lower the ethnic relevance of his candidacy for governor. Our data indicate that he succeeded in preventing the development of a large racially motivated opposition vote.

This outcome stands in striking contrast to the high level of racial voting in the recent mayoral elections in Chicago. That city has a long history of racial conflict. Tensions intensified in the years leading up to Harold Washington's election in 1983. The growth of the black population threatened the existing patterns of housing and schooling favored

[12] It is possible that antiblack voters anxious to conceal racist feelings would say that they were undecided rather than either express open support for Deukmejian and Honig or make the dissonant statement that they would vote for the black candidate. We were able to explore this possibility with the CP data. In that survey, voters were first asked which candidate they intended to vote for without the candidates being named. Respondents who said they did not know how they would vote were then read the names of the candidates and asked whom they would support "if the election were held today." The proportion of white respondents who failed to state a voting intention on the initial inquiry ranged from 27.5% of those sampled in the case of the gubernatorial contest to 61.3% and 54.2% in the McCarthy–Hallett and Riles–Honig elections respectively.

The initial group of "undecided" or "don't know" respondents were distinguished mainly by their lower level of formal education. They were not, as a group, more likely to give antiblack answers on our measures of racial attitudes. Nor did racial attitudes have a stronger impact on their voting decisions than on the choices of those willing to state an initial preference. As a group, undecided voters tended to fall in between Bradley and Deukmejian supporters on the various predictors of voting intention in our model. There thus is no compelling evidence that the self-proclaimed undecided white respondents were covert racists. The most likely explanations for their reluctance to state a preference are genuine uncertainty and a simple lack of information about the candidates.

by white "ethnics." In addition, the city's fiscal crisis and cuts in federal spending during the Reagan administration mobilized the black community to protest conditions in the areas of welfare, health care, education, and political representation (Alkalimat and Gills, 1986). Washington's candidacy was the climax of several years of black mobilization. His campaign called for the empowerment of blacks and the economically disadvantaged, culminating in the slogan "It's our turn," and his assertive style made the challenge to the existing political order unmistakable.

Despite the official condemnation of politicians and editorial writers, racist literature was widely circulated by groups opposed to Washington (Alkalimat and Gills, 1984:135). Indeed, Baker and Kleppner (1986) call the 1983 Chicago election a "race war" and Alkalimat and Gills (1984:137) characterize it as a conflict between black power and racism. The mobilization of black voters enabled Washington to win despite the massive defection of white Democrats to the Republican Bernard Epton, who had become their "hope" simply because of his race.

In 1989, as Bradley's own political fortunes waned, black candidates for mayor in New York and New Haven and for governor in Virginia emulated his "color-blind" campaign style to lessen the impact of their race on white voters. However, the 1983 mayoral election in Philadelphia demonstrates that this strategy does not always work. Wilson Goode's personal style also was managerial rather than agitational, and he too avoided appeals to racial solidarity in his campaign (Keiser, 1987). Nevertheless, the pattern of voting in Philadelphia resembled the mayoral contest in Chicago rather than the California gubernatorial election. It appears that when the black population is large, racial divisions are well entrenched, and the white candidate, like Frank Rizzo in Philadelphia, is an established symbol of hostility to blacks, the fact that the black candidate is moderate and conciliatory is not enough to diminish racial voting.

Additional research on biracial elections should specify more precisely the ways situational factors shape the reactions of white voters to black politicians.[13] The present analysis suggests some of the conditions under which race is not a critical factor in how a black candidate for a prominent office is evaluated. Tom Bradley's persistent effort to appear racially impartial led voters to evaluate him more as a liberal and a Democrat than as a black. On the other hand, when Harold Washington made the balance of power between the races a salient issue, race rather than party affiliation or political philosophy became the dominant influence on voting decisions. The strategic problem for a black candidate, of course, is how to frame a campaign that maximizes the benefits of black mobilization while minimizing the risk of white countermobilization. As long as the American public remains deeply divided on racial matters, this dilemma is likely to endure.

---

[13] Bradley ran for governor in a period of Republican and conservative gains in California. In 1982 he had the advantage of running against a less well known candidate. In his 1986 rematch with Governor Deukmejian, Bradley suffered an overwhelming defeat, winning only 33% of the white vote as compared to 44% in 1982 (Field Institute, 1986). This drop in support for him, however, was due to the popularity of his incumbent opponent rather than to an increase in racially motivated voting (Pettigrew and Alston, 1988). References to race were even less frequent during the campaign in 1986 than when Bradley ran in 1982. In fact, pollsters so completely discounted the role of race that they did not include any questions about racial issues in their preelection surveys, making it impossible to replicate our analysis for the 1986 election.

References

Alkalimat, A. and D. Gills (1984)
"Chicago." Pp. 53–179 in R. Bush (ed.) The New Black Vote. San Francisco: Synthesis Publications.

Becker, J. F. and E. E. Heaton, Jr. (1967)
"The election of Senator Edward W. Brooke." Public Opinion Quarterly 31: 346–358.

Baker, S. and P. Kleppner (1986)
"Race war Chicago style: The election of a black mayor, 1983." Pp. 215–238 in T. N. Clark (ed.) Research in Urban Policy, vol. 2. Greenwich, CT: JAI Press.

Bobo, L. (1983)
"Whites' opposition to busing: Symbolic racism or realistic group conflict?" Journal of Personality and Social Psychology 45: 1196–1210.

Boyd, R. (1972)
"Popular control of public policy: A normal vote analysis of the 1968 election." American Political Science Review 66: 429–449.

Campbell, A., P. E. Converse, W. Miller, and D. Stokes (1966)
Elections and the Political Order. New York: Wiley.

Carmines, E. and J. A. Stimson (1980)
"The racial reorientation in American politics." Pp. 199–218 in J. Pierce and J. Sullivan (eds.) The Electorate Reconsidered. Beverly Hills, CA: Sage.

Citrin, J. and D. P. Green (1990)
"The self-interest motive in American public opinion." Pp. 1–28 in S. Long (ed.) Research in Micropolitics, vol. 3. Greenwich, CT: JAI Press.

Field Institute (1983)
California Opinion Index. San Francisco: Field Institute. February.

——— (1986)
California Opinion Index. San Francisco: Field Institute. December.

Hahn, H., D. Klingman, and H. Pachon (1976)
"Cleavages, coalitions, and the black politician: The Los Angeles mayoralty elections of 1969 and 1973." Western Political Quarterly 29: 507–520.

Henry, C. (1983)
"The role of race in the Bradley-Deukmejian campaign." Pp. 41–53 in Critical Perspectives of Third World America, vol. 1. Berkeley: Ethnic Studies Student Union and Editorial Board Press.

Katz, P. (1976)
"Racism and social science: Towards a new commitment." Pp. 3–18 in P. Katz (ed.) Toward the Elimination of Racism. New York: Pergamon Press.

Keiser, R. (1987)
"Philadelphia." Unpublished manuscript, Department of Political Science, University of California, Berkeley.
Key, V. O., Jr. (1949)
Southern Politics in State and Nation. New York: Alfred A. Knopf.
Kinder, D. (1986)
"The continuing American dilemma: White resistance to racial change forty years after Myrdal." Journal of Social Issues 42:151–171.
Kinder, D., and D. O. Sears (1981)
"Prejudice and politics: Symbolic racism versus racial threats to the good life." Journal of Personality and Social Psychology 40:414–431.
McConahay, J. B. (1986)
"Modern racism, ambivalence, and the modern racism scale." Pp. 91–125 in J. F. Dovidio and S. L. Gaertner (eds.), Prejudice, Discrimination, and Racism. Orlando, FL: Academic Press.
McConahay, J. B., and J. C. Hough, Jr. (1976)
"Symbolic racism." Journal of Social Issues 32:23–45.
Pettigrew, T. (1976)
"Black mayoral campaigns." Pp. 14–29 in H. J. Bryce (ed.), Urban Governance and Minorities. New York: Praeger.
Pettigrew, T., and D. Alston (1988)
Tom Bradley's Campaigns for Governor: The Dilemma of Race and Political Strategies. Washington: Joint Center for Political Studies.
Schuman, H., and L. Bobo (1988)
"Survey-based experiments on white racial attitudes toward racial integration." American Journal of Sociology 94:273–299.
Schuman, H., C. Steeh, and L. Bobo (1985)
Racial Attitudes in America: Trends and Interpretations. Cambridge: Harvard University Press.
Sears, D. O. (1988)
"Symbolic racism." Pp. 53–84 in P. A. Katz and D. A. Taylor (eds.), Eliminating Racism: Profiles in Controversy. New York: Plenum.
Sears, D. O., and J. Citrin (1985)
Tax Revolt. Enl. ed. Cambridge: Harvard University Press.
Sears, D. O., J. Citrin, and R. Kosterman (1987)
"Jesse Jackson and the southern white electorate in 1984." Pp. 209–225 in L. W. Moreland, R. P. Steed, and T. A. Baker (eds.), Blacks in Southern Politics. New York: Praeger.
Sears, D. O., and C. L. Funk (1988)
"Self-interest and American public opinion." Unpublished manuscript, UCLA Department of Psychology.
Sears, D. O., C. P. Hensler, and L. K. Speer (1979)
"Whites' opposition to 'busing': Self-interest or symbolic politics?" American Political Science Review 73:369–384.
Sears, D. O., L. Huddy, and L. G. Shaffer (1986)
"A schematic variant of symbolic politics theory, as applied to racial and gender equality." Pp. 159–202 in R. Lau and D. O. Sears (eds.), Political Cognition. Hillsdale, NJ: Erlbaum.
Sears, D. O., R. Lau, T. Tyler, and H. M. Allen, Jr. (1980)
"Self-interest versus symbolic politics in policy attitudes and presidential voting." American Political Science Review 74:670–684.
Sniderman, P. M., and P. Tetlock (1986)
"Symbolic racism: Problems of motive attribution in policy analysis." Journal of Social Issues 42:129–150.
Wolfinger, R. (1974)
The Politics of Progress. Englewood Cliffs, NJ: Prentice-Hall.

THE EFFECTS OF PRESTIGE NAMES IN QUESTION WORDING

ERIC R. A. N. SMITH AND
PEVERILL SQUIRE

Abstract Although survey researchers are often warned against using prestige names in questions (e.g., identifying Contra aid as "President Reagan's policy"), prestige names are still commonly used. To learn more about the effects of using prestige names, we analyze two sets of experiments—on judicial confirmation elections and on an initiative on tax indexing. The results indicate that prestige names do more than shift responses in one direction—they eliminate the effect of education on DKs and provide more of a political basis for the responses.

A well-known rule for survey researchers writing questions is, Do not use "prestige names" in questions. That is, one should avoid the use of names of people or organizations (e.g., President Reagan, the Supreme Court) that add prestige to one side of a question unless, of course, the question is about those persons or organizations. So if one is interested in asking about, say, whether the United States should support the Contras in their effort to overthrow the government in Nicaragua, one should not refer to it as "President Reagan's policy." The information that the president takes one side of the issue may affect the responses by giving people an additional cue for answering the question. Instead of answering the question based on their beliefs about the situation in Central America, people may answer based on their evaluations of the president.

The usual treatment of this problem in the question wording literature is to explain that prestige names may bias questions and to give one or two examples of differently worded questions in which the marginal distributions have changed (Rugg and Cantril, 1944; Lorge, 1936; Stagner, 1941; Sudman and Bradburn, 1982). Analysis rarely goes beyond marginal distributions. Because adding prestige names to

ERIC R. A. N. SMITH is Assistant Professor of Political Science at the University of California at Santa Barbara. PEVERILL SQUIRE is Associate Professor of Political Science at the University of Iowa.

Public Opinion Quarterly Volume 54:97–116 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

opinion questions results in confusing the opinion with the evaluation of the person or symbol named, researchers do not use such questions very often, and those investigating question wording effects have largely ignored them in recent years (the exceptions are Rosen, 1973; Sigelman and Thomas, 1984; and Tarantino and Jednak, 1972).

In this article, we will explore what difference prestige names make to respondents using two different sets of questions asked by the California Poll in 1982. Instead of merely showing that the use of prestige names does make a difference, we will look at what sorts of people it affects, how it affects them, and what information is contained in prestige names.

This is important for several reasons. First, major polls still occasionally ask questions using prestige names to identify policy positions. For instance, ABC News asked the following question in March 1986:

> President Reagan is asking Congress for new military aid for the Nicaraguan rebels known as the "Contras." Do you agree or disagree with Reagan that Congress should approve the money? (Public Opinion, 1986:29)[1]

ABC is not alone; prestige names continue to be used in survey questions.[2] This practice will probably continue because, as Sudman and Bradburn (1982:135) point out, pollsters can defend it on the grounds that "certain issues are so closely related to individuals or organizations that it is unrealistic to separate them."[3]

Second, the practice of using prestige names in questions was more common in the past, and researchers interested in analyzing older data may benefit from a better understanding of how prestige names influence responses. Third, from the point of view of practical politics, we can learn what happens when obscure issues or candidates are associated with well-known names.

Because analysis of prestige names has been restricted to comparing marginals, little is known about the effects of prestige names. Three points stand out. First, it is not at all clear that adding a prestige name regularly—or even usually—helps the side associated with the name. In an experiment reported by Rugg and Cantril (1944), associating Franklin Roosevelt's name with a proposal to send Sumner Welles to

---

[1] A month earlier, a substantively similar question which did not invoke the name of President Reagan was asked by the CBS News/*New York Times* poll (Public Opinion, 1986:29). Nine percent fewer of the respondents supported aid for the Contras (25% vs. 34%). The difference may have other causes, but it does make one suspect that ABC's question exaggerated the Contras' support.

[2] The issue of *Public Opinion* from which the Contra example was taken, for instance, includes three other items with prestige names (pp. 27, 35).

[3] Sudman and Bradburn, however, maintain that it is a poor practice. They say it is far better to separate issues from sources if at all possible.

Europe reduced support among those who disapproved of Roosevelt, but it did not increase support among those who approved of him. In a pair of classroom experiments in Iowa and Kentucky, Sigelman and Thomas (1984) found that attributing the (nonexistent) Agricultural Trade Act to President Reagan or Iowa Governor Robert Ray increased support for the act in Iowa, but attributing the act to Reagan or Kentucky Governor John Y. Brown decreased support in Kentucky.

Second, it is not the case that prestige names always influence the marginal distributions of questions. Rugg and Cantrill (1944:40) reported that on a number of foreign policy items (e.g., the Lend-Lease bill) the addition of President Roosevelt's name made no difference at all. Similarly, Tarantino and Jednak (1972) found that attributing the quotation, "a little rebellion now and then is a good thing" to a variety of people from Thomas Jefferson to Lenin and Mao made no difference.

Third, prestige names usually seem to provide information to and influence those who have no other basis on which to answer the questions. The evidence for this is that most published examples of prestige names show that adding a name reduces the proportion of those without opinions. However, Rugg and Cantril (1944:39) present an example of a salient issue (whether the United States should do more or less to help Britain in June 1941) on which opinion changed substantially but the proportion of DKs remained the same. Thus the findings about the effects of prestige names are somewhat contradictory.

The First Experiment

Two experiments on prestige names in survey questions were conducted by the California Poll in 1982—the first on state supreme court confirmation elections and the second on a tax-indexing proposition on the California ballot.[4] We will begin with the state supreme court confirmations.

In the November general election of 1982, four justices were on the ballot for confirmation. That is, voters had to decide whether to keep them in office or remove them; there were no opposing candidates. In the August 1982 California Poll, a split ballot experiment was conducted in which respondents were asked one of two versions of a vote intent question about each of the four justices. The first version (N =

---

[4] These experiments were conceived and carried out by the Field Research Corporation. The data sets, California Polls 8202 and 8205, were provided by the University of California State Data Program, Berkeley. Neither of these organizations is responsible for the analysis or interpretation of the data appearing here.

408) asked whether the respondent would vote to confirm or remove the justice. The second version ($N = 402$) asked whether the respondent would vote to confirm or remove and stated which governor—the incumbent governor, Democrat Jerry Brown, or the former governor, Republican Ronald Reagan—had appointed the justice. The first question was:

> In the November election voters will be asked to vote for or against a number of State Supreme Court Justices who are now serving in office. As I read each one please tell me as of now whether you would vote to have that person remain on the court or be removed. The first is [READ NAME OF PERSON CHECKED]. Would you vote to have him stay on the court or be removed?

The second question was the same as the first except for the addition of a clause identifying which governor appointed the justice. The questions under this format read, for example, ". . . Justice Cruz Reynoso, appointed by Governor Jerry Brown," or ". . . Justice Frank Richardson, appointed by Ronald Reagan when he was governor." Thus, the second questions added prestige names.

To put the experiment in context, we should say something about nonpartisan judicial elections. Most voters in judicial elections know very little about the candidates (Johnson, Shaefer, and McKnight, 1978; Lovrich and Sheldon, 1983). Lack of knowledge certainly seems to characterize the California elections. When asked by the California Poll in October 1982 to name the chief justice of the state supreme court, only 43% of the respondents correctly identified Rose Bird, despite the fact that she was very controversial (Stolz, 1981). This suggests that her less-controversial colleagues were even less well known.

Another factor in this election was that the justices differed on racial, ethnic, and political lines. Three of the justices were appointed by Jerry Brown: Cruz Reynoso, the first Hispanic to serve on the court; Allen Broussard, the first black justice on the court to come before the voters; and Otto Kaus. Reynoso and Broussard were considered liberal, and Kaus was a moderate. The Reagan appointee was Frank Richardson, a white conservative.

Analysis

The marginal distributions for our eight questions are presented in Table 1. These are the data typically presented in the analysis of prestige names. The first point to note is that the addition of the identity of the appointing governor sharply reduced the number of respondents

Table 1. Support for Judicial Confirmation, With and Without Information

Justice (Governor)[a]	Confirm	Remove	No Opinion
Reynoso	31%	11	58
Reynoso (Brown)	34%	27	39
Kaus	29%	7	63
Kaus (Brown)	33%	21	45
Broussard	31%	8	61
Broussard (Brown)	32%	22	46
Richardson	36%	5	59
Richardson (Reagan)	41%	15	44

SOURCE: Data collected by the Field Research Corporation, California Poll 8205, 23–27 August 1982.

NOTES: N (no information) = 408; N (information) = 402.

[a] The governor's name appears in parentheses for the questions in which the appointing governor was identified.

having no opinion. The drop ranges from 15% to 19%. This decrease in DKs fits the predominant pattern of previous experiments.

The second point to note is that the largest increases were among the percentages saying that they would remove the justices. While there were only 1%–5% increases in the number saying that they would vote to confirm, there were 10%–16% increases in the number saying that they would vote to remove. We do not know, however, whether adding the names simply moved some respondents from the DK category to holding opinions or moved some respondents back and forth between the confirm and remove options as well.

To find out more about how respondents reacted to the extra information, we estimated two sets of probit equations. The first set of equations is designed to explain whether people have opinions; the second set is intended to explain the confirm/remove choice. Because both dependent variables are dichotomies, we use probit rather than simple regression (Hanushek and Jackson, 1977).

Several variables might be expected to influence whether a respondent has an opinion.[5] Education should certainly cause opinion holding (Campbell et al., 1960; Schuman and Presser, 1981) with the better educated being more likely to have opinions. We expect strength of party identification to be important because stronger partisans have been shown to know more about politics than weak partisans or independents (Campbell et al., 1960; Keith et al., 1977). Similarly, party identification might influence opinion holding because Republicans have been found to be more knowledgeable than Democrats about judicial candidates (Lovrich and Sheldon, 1983). Evaluation of the appointing governor may operate in a manner similar to party identification. Ideological identification could matter because conservatives may know more about politics than liberals. Finally, being black or Hispanic may have an independent effect on opinion holding. Certainly Hispanics, given the ease with which Cruz Reynoso's name can be identified, would be expected to express opinions. Blacks may not be as aware of Broussard's race, but some effect may be found.

---

[5] Variable coding is given in the Appendix. A more complete description of these variables and their expected relationships is contained in Squire and Smith, 1988.

Most of these variables can also be hypothesized to affect what opinion is held. We expect party identification to have some influence because three of the justices were appointed by a Democrat and the other one by a Republican. The rating of the appointing governor should also be important, with those liking the governor supporting his appointees and those not liking the governor opposing them. Likewise conservatives should be likely to support Reagan's selections and oppose Brown's, with liberals behaving in the opposite fashion. Because race can substitute for party labels in nonpartisan elections (Lovrich, Sheldon, and Wassman, 1988; Pomper, 1966), Hispanics should be expected to back Reynoso and blacks should be expected to support Broussard. Finally, more-educated people may be more supportive of an independent judiciary and thus be more likely to favor confirmation for each justice.

The probit equations on opinion holding are presented in Table 2. Two equations are presented for each justice; the first explains opinion holding for the question without the appointer's name, and the second explains opinion holding for the question with the name of the appointing governor.

The most striking finding in these equations involves the role of education. As expected, education has a large impact on opinion holding in the no-information questions. In two of the equations it is significant at the usual .05 level; in the other two—Kaus and Richardson—it just misses at $p < .053$ and $p < .063$. The better educated the respondent is, the more likely he or she is to take a position on confirmation. The addition of the identity of the appointing governor, however, removes this relationship. In all four cases, the extra information reduces the education coefficients to substantive and statistical insignificance. These changes in education coefficients between the two versions of the questions are not all statistically significant—the shifts in the Reynoso and Kaus equations are borderline at $p < .08$, and

Table 2. Probit Equations With and Without Prestige Name: Holding an Opinion or Not as Dependent Variable

	Reynoso		Broussard		Kaus		Richardson	
Variable	No Name	Name	No Name	Name	No Name	Name	No Name	Name
Education	.204**	.054	.148**	.049	.117*	−.036	.111*	−.001
	(.061)	(.060)	(.060)	(.059)	(.060)	(−.058)	(.060)	(.058)
Strength of party ID	.099	.059	.186**	.128*	.098	.060	.239**	.037
	(.075)	(.073)	(.075)	(.072)	(.075)	(.072)	(.075)	(.072)
Party ID	−.096**	−.093**	.002	−.074*	−.074	−.106**	−.026	−.076*
	(−.040)	(−.041)	(.039)	(−.040)	(−.039)	(−.040)	(−.039)	(−.040)
Ideology	.013	.005	.021	.011	.016	−.017	.013	−.017
	(.043)	(.045)	(.043)	(.044)	(.043)	(−.044)	(.042)	(.044)
Reagan rating	−.130**	−.146**	−.012	−.094	−.036	−.172**	−.023	−.086
	(−.072)	(−.068)	(−.072)	(−.067)	(−.071)	(−.067)	(−.071)	(−.066)
Brown rating	−.143**	−.111	−.160**	−.157**	−.027	−.098	−.047	−.123*
	(−.069)	(−.067)	(−.068)	(−.066)	(−.068)	(−.066)	(−.067)	(−.066)
Hispanic	.798**	.697**	.516**	.229	.282	.164	.274	.149
	(.242)	(.267)	(.234)	(.242)	(.245)	(.242)	(.231)	(.241)
Black	.460	.110	.309	.298	.278	.196	.209	.244
	(.247)	(.262)	(.245)	(.262)	(.245)	(.262)	(.243)	(.262)
Constant	−.033	1.027**	−.936**	.542	−.563	1.346**	−1.023**	1.015**
	(−.474)	(.487)	(−.474)	(.476)	(−.473)	(.481)	(−.472)	(.478)
N	389	384	389	384	389	384	389	384

SOURCE: Data collected by the Field Research Corporation, California Poll 8205, 23–27 August 1982.
NOTE: The column entries are probit coefficients. The numbers in parentheses are standard errors.
* $p < .075$. ** $p < .05$.

Table 3. Probability of Holding an Opinion, by Education

Justice	Less Than High School	High School Grad	Some College	College Grad	Grad +	Diff Grad–HS
Reynoso	29	36	44	52	61	32
Reynoso (Brown)	57	59	61	64	66	9
Broussard	30	35	41	46	52	22
Broussard (Brown)	51	53	55	57	59	8
Kaus	29	34	38	43	48	19
Kaus (Brown)	57	56	55	54	53	−4
Richardson	34	39	43	47	52	18
Richardson (Reagan)	56	57	57	57	57	1

SOURCE: Data collected by the Field Research Corporation, California Poll 8205, 23–27 August 1982.

NOTE: Cell entries are predicted probabilities of holding opinions.

the Richardson and Broussard shifts are even weaker at $p < .18$ and $p < .25$ respectively. Still, the fact that all the changes are substantively large and in the same direction makes the results important.

Because probit coefficients are inherently somewhat difficult to interpret, Table 3 shows the effects of education with and without information.[6] The two rows on Reynoso provide a good example. The first row shows that when voters are not told who appointed Reynoso, if all voters had less than a high school education, only 29% would give opinions. At the other extreme, if all voters had attended at least some graduate school, 61% would express opinions—a 32% increase over the less-than-high-school possibility. When the voters are told that Governor Brown appointed Justice Reynoso, the difference between the least and most educated is only a statistically insignificant 9%.

This finding indicates that the least educated are responding most strongly to the additional information. This is particularly interesting since every study of DK responses indicates that education is a major cause of opinion holding; the least educated are the most likely not to have opinions (Bartels, 1986; Bishop, Oldendick, and Tuchfarber, 1980; Converse, 1976; Ferber, 1966; Francis and Busch, 1975; Schuman and Presser, 1981). Here we see that adding the prestige names nullifies that result. The advantage of the well educated is eliminated.

---

[6] The probabilities in Table 3 are calculated from the equations in Table 2 using a method explained in Wolfinger and Rosenstone, 1980:123.

Returning to Table 2, several other variables influence opinion holding in the no-information equations. As expected, Hispanics are more likely to have opinions on Reynoso's confirmation. Blacks show a similar relationship, but it does not quite reach statistical significance. For Broussard, Hispanics are surprisingly still more likely to have opinions than blacks. We cannot find any ready explanation for this finding; race and ethnicity have no effects on holding opinions about either of the other two justices.

Of the explicitly partisan variables—party identification and ratings of Brown and Reagan—all three influence opinion holding about Reynoso. Only one of these variables, the Brown rating, influences opinion about Broussard, and none of these variables influences opinion on either Kaus or Richardson.

When the questions identify the governor who appointed each justice, there are several changes. As noted above, education no longer matters. Being Hispanic continues to work in the Reynoso equation, but being Hispanic or black is not significant in any other equation. Instead, partisanship seems to have more influence in the process. We must be tentative in this claim because, although the coefficients become significantly different from zero when the appointing governor's name is given (at $p < .075$ or better), none of the coefficients is significantly greater than its counterpart in the no-information equations. Nevertheless, all four equations show Republicans becoming more likely to state opinions, and at least two of the changes are large enough to be considered substantively important. We believe that together they form a general pattern of increasing importance for party identification. In sum, the important result in Table 2 is that when the identity of the appointing governor is given, education, race, and ethnicity cease to make any difference (except in the case of Hispanics and Reynoso) and party identification begins to have an influence.

We now turn to Table 4 to see what causes the content of these opinions. Overall, very little structures the opinions without the identity of the appointing governor. Only two of the 28 independent variables in the four no-information questions are statistically significant at the .05 level. Evaluations of Governor Brown are positively related to favoring confirmation for two of his appointees, Broussard and Reynoso. No other variables are significant.

The addition of the governors' names changes the picture. In all four information equations, evaluation of the appointing governor is strongly related to support for confirmation. In the Reynoso and Broussard equations, where the Brown rating was already significant, the changes are trivial. But in the Kaus and Richardson equations there are large and significant increases in the sizes of the Brown and Reagan ratings. In addition, in two of the equations (Broussard and Reynoso)

Table 4. Probit Equations With and Without Prestige Name: Position on Confirmation as Dependent Variable

Variable	Reynoso		Broussard		Kaus		Richardson	
	No Name	Name	No Name	Name	No Name	Name	No Name	Name
Education	.169	.172	.059	.130	.210	.016	.158	.082
	(.109)	(.089)	(.107)	(.091)	(.116)	(.093)	(.118)	(.092)
Party ID	.122	.199*	−.042	.131*	.107	.042	−.009	−.024
	(.077)	(.058)	(.077)	(.058)	(.079)	(.062)	(−.073)	(−.063)
Ideology	−.014	−.019	−.022	.070	.116	.068	.046	−.110
	(−.080)	(−.063)	(−.079)	(.064)	(.089)	(.069)	(.081)	(−.060)
Reagan rating	−.180	−.066	−.165	−.094	−.025	−.104	−.032	.458*
	(−.143)	(−.096)	(−.139)	(−.098)	(−.161)	(−.101)	(−.145)	(.101)
Brown rating	.423*	.630*	.467*	.419*	.226	.614*	.130	.037
	(.134)	(.105)	(.142)	(.106)	(.141)	(.115)	(.139)	(.104)
Hispanic	.293	.034	.239	−.135	.363	.289	.038	.324
	(.397)	(.331)	(.405)	(−.376)	(.465)	(.447)	(.435)	(.374)
Black	−.375	2.913	−.580	2.832	.240	2.723	−.373	.730
	(.484)	(4.760)	(−.445)	(5.345)	(.610)	(4.972)	(−.438)	(.375)
Constant	−.546	−2.951*	.360	−1.604*	−.973	−1.361*	.376	−.475
	(−.830)	(.637)	(.799)	(−.621)	(−.924)	(.628)	(.845)	(−.594)
N	171	236	159	211	150	213	168	218

SOURCE: Data collected by the Field Research Corporation, California Poll 8205, 23–27 August 1982.
NOTE: The column entries are probit coefficients. The numbers in parentheses are standard errors.
* $p < .05$.

party identification is strong and significant, although only the increase in the Broussard coefficient is significant ($p < .075$). The more Democratic the respondent, the more likely he or she is to favor confirmation of these appointees of the Democratic governor. In none of the equations is education, ideology, race, or ethnicity significant.

The overall effect of adding the prestige names to these questions is quite clear. The content of the no-information answers had no structure that we could discover. The two significant coefficients, ratings of Governor Brown, were presumably specific to these questions, since the Governor had only recently appointed the first black and first Hispanic to the state supreme court. Thus, we are left with a weak explanation of two of the confirmation votes and no explanation of the other two confirmation votes. The questions with information were structured along partisan lines; party identification and evaluations of the appointing governors (which partly reflect party identification) were the key influences on both opinion holding and content of opinion.

The Second Experiment

In the June 1982 primary election, voters were faced with an initiative, sponsored by Howard Jarvis, that would require state personal income taxes to be indexed for inflation. This was the fourth round of the California tax revolt (Sears and Citrin, 1982), which had started in 1978 with the passage of Proposition 13, sponsored by Howard Jarvis and Paul Gann. Proposition 7 in 1982 was the last successful effort by the tax revolt forces.

Because of the success of the tax revolt and the enormous impact it had on California and the nation, Jarvis and Gann became celebrities. Jarvis, the more colorful of the two, was the subject of cover stories in both *Time* and *Newsweek*. By 1980, however, Jarvis's abrasive manner and bullying, insulting style of attacking his opponents had caused him image problems (Sears and Citrin, 1982:29–39). Thus in 1982 Jarvis was a well-known figure about whom many voters had strong opinions.

In the March 1982 California Poll, a split-ballot experiment was conducted in which respondents were asked one of two sets of questions about Proposition 7. Respondents were first asked whether they had heard of the proposition. If they had heard of it, they were asked how they intended to vote. Then regardless of whether they had heard of it or not, it was described to them and they were again asked how they intended to vote. In one version of the questions ($N = 614$), respondents were told that the proposition was sponsored by Howard Jarvis; in the other version ($N = 603$), they were not told. The first version of

Table 5. Support for Tax Indexing Proposition With and Without Prestige Names

	Yes	No	No Opinion	*N*
Heard about Prop 7	15%	85		603
Heard about Prop 7 (Jarvis)	27%	73		614
If heard				
How vote	60%	13	26	91
How vote (Jarvis)	52%	23	25	167
Prop 7 described				
How vote	53%	27	20	603
How vote (Jarvis)	43%	42	15	614

SOURCE: Data collected by the Field Research Corporation, California Poll 8202, 18–23 March 1982.

the questions were:

> Have you seen or heard anything about a constitutional amendment, Proposition 7, sponsored by Howard Jarvis, that will be on the next statewide election, having to do with the indexing of state personal income taxes?[7]
>
> [IF YES HAVE HEARD] From what you have seen or heard, would you vote yes or no on Proposition 7, the measure to index state personal income taxes?
>
> Proposition 7, sponsored by Howard Jarvis, would require that personal tax brackets be adjusted to account for the full amount of inflation each year. It has been estimated that this change would reduce the amount of money the state takes in by about 230 million dollars the first year and larger amounts in later years. Suppose you were voting on this measure today. Would you now vote yes or no?

In the second version, the phrase "sponsored by Howard Jarvis" was dropped.

The marginal distributions for the questions are presented in Table 5. The basic patterns of the first experiment recur here. The number claiming to have heard of the proposition increases by 12% when Jarvis's name is associated with it (from 15% to 27%). Similarly, the number who have no opinions when the proposition is described falls by a statistically significant 5% (from 20% to 15%) when Jarvis's name is added to the question. Only for the second question among those who claimed that they had heard about it did the percentage of DKs not fall when Jarvis's name was added.

---

[7] There were two minor changes in the question wordings. In the version without Jarvis's name, the phrase "the next statewide election" was changed to "the June ballot." In the follow-up question for those who had heard of it, the phrase "to index state personal income taxes" was changed to "indexing state personal income taxes." While these trivial changes are unfortunate, we doubt that either one could affect the results of the experiment.

Also as in the first experiment, the balance of those favoring and opposing the proposition changed among both those who had heard of it and the full sample. Adding Jarvis's name to the questions caused an 8% decrease in support for the proposition among those who had heard of it and a 10% decrease in the full sample. The increases in opposition to the proposition were 10% and 15% among the two groups. Indeed, telling the full sample that Jarvis had sponsored the proposition changed the outcome from a landslide 53%–27% margin to a virtual dead heat at 43%–42%.

The shift in responses when Jarvis's name is added to the question is an important finding. The fact that there was a drop in the number of those favoring the Proposition indicates that adding Jarvis's name caused a change among those who answered the questions. This is not a case of DKs forming opinions and choosing sides, but a case of a prestige name causing people to give different substantive responses.

To examine the effects of Jarvis's name in more detail, we again turn to probit analysis. Specifically, we will examine three pairs of dependent variables—who claims to have heard of Proposition 7, who has an opinion on Proposition 7, and what that opinion is.

For our independent variables, we have turned to the literature on the tax revolt (Citrin, 1979; Sears and Citrin, 1982). Support for previous tax revolt initiatives came from people who were high income, high education, older, white, conservative, not government employees, residents of Southern California, and home owners. Because Proposition 13 was specifically aimed at property taxes and home owners, but tax indexing was not aimed at home owners, it seemed to us that home ownership was not very likely to cause support for Proposition 7. Indeed, our initial analysis indicated that home ownership had no effect on any of our dependent variables; therefore, we dropped it from further analysis.

We should also point out that here, unlike the first experiment, we see no reason to distinguish between blacks and Hispanics. While one justice was black and another Hispanic, Proposition 7 would be likely to affect the interests of all minorities equally. Thus we include only the variable white/nonwhite.

On the question of opinion holding, we expect that education, strength of party identification, and race might contribute to the explanation for the reasons we suggested in the first experiment. In addi-

Table 6. Probit Equations With and Without Prestige Name: Holding an Opinion or Not as Dependent Variable

	Heard of Prop 7		Opinion on Prop 7	
Variable	No Name	Name	No Name	Name
Age	.013*	.018*	−.014*	−.010*
	(.004)	(.004)	(−.004)	(−.004)
Govt employee	−.07	−.09	−.06	.22
	(−.18)	(−.15)	(−.17)	(.19)
Education	.14*	.11*	.10	−.01
	(.06)	(.05)	(.06)	(.06)
Race	.25	.02	.01	.01
	(.19)	(.15)	(.17)	(.17)
Income	.07*	.15*	−.02	.00
	(.03)	(.03)	(−.03)	(.03)
Region	.22	−.21	−.05	.43*
	(.14)	(−.13)	(−.13)	(.15)
Strength of party ID	.01	.08	−.12	.14*
	(.07)	(.07)	(.07)	(.07)
Constant	−2.67*	−2.69*	1.70*	.88*
	(−.38)	(−.33)	(.35)	(.33)
N	541	565	541	565

SOURCE: Data collected by the Field Research Corporation, California Poll 8202, 18–23 March 1982.

* $p < .05$.

tion to those variables, it might be the case that those who were the strongest proponents of the earlier tax revolt initiatives might be more likely to know about and hold opinions on the current tax revolt initiative. Therefore, we also included age, income, government employee status, and region.

The probit equations on knowledge and opinion holding are presented in Table 6. Here we are looking at two dependent variables. In the first two columns, the dependent variable is whether the respondent claimed to have heard of Proposition 7. In the third and fourth columns, the dependent variable is whether the respondent had an opinion on the Proposition after it was described in the third question.

Looking first at who had heard of the Proposition, we see that the addition of Jarvis's name made little difference. Education, income, and age all have significant effects in the expected directions in both equations. Moreover, the coefficients are all of roughly the same size. The only other variable that comes even close to significance is region.

In the first equation, among those who are not told that Jarvis is the sponsor, Southern Californians are more likely to have heard of the proposition ($p < .11$), while in the second equation, among those who were told about Jarvis's sponsorship, Southern Californians are less likely to have heard of it ($p < .09$). While neither coefficient is significantly different from zero, they are significantly different from one another. We had expected Southern Californians to be more likely to respond to Jarvis's name, but it turns out that Jarvis's name had a much greater effect outside of his home base in the south.

Beyond the details of the coefficients, we should focus on the central point of the experiment. Even though adding Jarvis's name to the question significantly increased the number saying that they had heard of the proposition, it seems to have had no effect on the relationships.

We now turn to the third and fourth columns of Table 6, which present the equations for opinion holding. Here we have some clear evidence that adding Jarvis's name to the question changed the results. In the third equation, without Jarvis's name, we see that only age has a significant effect. Contrary to our hypothesis, the younger respondents are more likely to hold opinions. A second unexpected finding is that education has no effect in this equation. When Jarvis's name is added to the question, both region and strength of party identification also become statistically significant in the expected directions—Southern Californians and strong party identifiers are more likely to hold opinions. Moreover, these changes are fairly large; in the no-name question, both coefficients were negative. In short, like the previous experiments on judicial confirmation, adding the prestige name altered the relationships between the independent variables and opinion holding.

We now turn to Table 7 to examine the cause of support for Proposition 7. The results are similar to those of the first experiment. When respondents are not told that Jarvis sponsored the proposition, we see that there is little by the way of structure. Only age and government employee status are significant; those who are older and those who are not government employees are more likely to support tax indexing.

Looking at the equation in which respondents were told that Jarvis was the proposition's sponsor, we see several changes. Age and government employee status are no longer significant (although only the change in age is significant; the shift in government employee status is trivial). Instead, race, region, and ideology are significant. However, while the changes in race and region are substantively large, they fail to be statistically significant ($p < .14$ and $p < .19$, respectively). These variables work in the expected direction. Whites, Southern Californians, and conservatives are more likely to support tax indexing; minorities, non–Southern Californians, and liberals are more likely to oppose it.

Table 7. Probit Equations With and Without Information: Position on Proposition as Dependent Variable

Variable	No Name	Name
Age	−.015*	−.004
	(−.004)	(.004)
Govt employee	−.36*	−.29
	(−.17)	(−.17)
Education	.00	.02
	(.07)	(.06)
Race	.07	.45*
	(.18)	(.16)
Income	.06	.05
	(.03)	(.03)
Region	.05	.29*
	(.15)	(.14)
Ideology	−.06	−.09*
	(.04)	(.03)
Constant	.97*	−.20
	(.34)	(−.34)
N	361	382

SOURCE: Data collected by the Field Research Corporation, California Poll 8202, 18–23 March 1982.
* $p < .05$.

In sum, without Jarvis's name in the question, responses are based on age and government employee status; with Jarvis's name, the effect of age is eliminated, and the effects of race and region increase. Thus adding Jarvis's name to the question changed the pattern of responses to the questions.

Discussion

In thinking about the results of our two experiments, an important difference between the experiments must be taken into account. In the first experiment, the subject—judicial confirmations—was one about which few voters were likely to know much. Moreover, aside from the name of the appointing governors, the questions did not provide any information. In the second experiment, the subject was one about which voters might know something—tax indexing. Moreover, both versions of the question contained a brief explanation of the provisions of the proposition. Thus we may presume that the level of respondent information was much higher in the second experiment than in the first.

Let us consider our experiments, beginning with the results about opinion holding. Two general conclusions seem to follow from these findings. First, the effect of education on opinion holding is greatest when respondents have the least information, and it decreases as respondents know more. We draw this inference because (1) the effect of education was greatest in our judicial confirmation experiment when respondents were not told who appointed each justice (i.e., when they had no information) and (2) in the first experiment when respondents were told who appointed the justices and in the second experiment when the proposition was explained in the questions (i.e., when respondents did have information), education had no effect on opinion holding.

Second, adding prestige names changes the sources of opinion holding, and changes them in a way that is not likely to be politically neutral. In the first experiment, the new source of opinion holding was party identification. In the second experiment, one of the new sources (strength of party identification) was partisan and the other (region) was not. Yet given the greater conservatism and the greater popularity of Howard Jarvis in Southern California, the regional change in responses was likely to have a political impact. Thus in both experiments, adding the prestige names to the questions produced a new political mix among those who offered opinions. Of course, other questions about other issues may yield increases in opinion holding that are politically neutral, but there is certainly no reason to expect that to happen. In general, we should expect prestige names to alter the political balance of those with opinions; the only question is how.

Let us now turn to the question of what opinions people had. Here the conclusion we draw is that using prestige names adds a partisan component to the questions, but that the partisan component seems to be greatest when respondents know the least. We base this conclusion on the pattern of effects of the partisan variables in the two experiments.

In the first experiment, ratings of Governor Brown were significant in two of the equations without the name of the appointing governor; nothing was significant in the other two equations. When the governors' names were added, the name of the appointing governor was significant in all four equations, and party identification became more important in two of the equations. That is, the results became more partisan. In the second experiment, age and government employee status had effects without Jarvis's name, but age ceased to have any effect, and race and region had larger effects, when his name was mentioned. As we noted above, race and region do have partisan impacts, but not as much as variables such as party identification and ratings of governors. Thus in the second experiment, when people had more knowledge, the political impact of the prestige name was smaller than in the first experiment.

In general, prestige names are additional stimuli in the questions and thus are additional sources of variance to be explained. Adding prestige names does not simply shift all respondents in one direction. Instead, some become more favorable and others less so as partisanship and other factors begin to structure responses. Given our findings about the partisan basis of holding opinions on issues when prestige names were included in the questions, this result should have been expected.

Our data imply that the interpretation of questions with prestige names is difficult because respondents choose sides not only on the basis of the content of the issues, but on the basis of the names associated with the different positions as well. Thus the results of the ABC Contra aid question quoted above, for instance, should not be interpreted as indicating support for Contra aid, but as indicating some unknown mixture of support for Contra aid and approval of President Reagan. Therefore, when issues can be separated from the names of supporters or opponents, the conventional wisdom—that prestige names do not belong in questions—yields far more easily interpreted results. When a researcher decides that the two cannot be separated—and there may be situations when that is the case—then the burden falls on the researcher to be cognizant of the complications inherent in interpreting the results.

Appendix: Coding of Variables

Age: actual age in years
Black: 0 = nonblack; 1 = black
Brown Rating: 1 = very poor; 2 = poor; 3 = fair; 4 = good; 5 = excellent
Education: 1 = less than high school; 2 = high school or vocational school graduate; 3 = some college; 4 = college graduate; 5 = some graduate education
Government Employee: 0 = no government employee in household; 1 = government employee in household
Hispanic: 0 = non-Hispanic; 1 = Hispanic
Ideology: 1 = strong conservative; 2 = weak conservative; 3 = leaning conservative; 4 = middle of the road; 5 = leaning liberal; 6 = weak liberal; 7 = strong liberal
Income: 1 = \$0–\$7,000; 2 = \$7,000–\$14,999; 3 = \$15,000–\$24,999; 4 = \$25,000–\$39,999; 5 = \$40,000 or more
Party Identification: 1 = strong Republican; 2 = weak Republican; 3 = lean-

ing Republican; 4 = independent; 5 = leaning Democrat; 6 = weak Democrat; 7 = strong Democrat
Race: 0 = nonwhite; 1 = white
Reagan Rating: 1 = very poor; 2 = poor; 3 = fair; 4 = good; 5 = excellent
Region: 1 = Los Angeles or Orange Counties; 0 = other county
Strength of Party Identification: 1 = independent; 2 = leaner; 3 = weak partisan; 4 = strong partisan

References

Bartels, Larry M. (1986)
"Issue voting under uncertainty: An empirical test." American Journal of Political Science 30:709–728.
Bishop, George, Robert Oldendick, and Alfred Tuchfarber (1980)
"Experiments in filtering opinions." Political Behavior 2:339–370.
Campbell, Angus, Philip Converse, Warren Miller, and Donald Stokes (1960)
The American Voter. New York: John Wiley.
Citrin, Jack (1979)
"Do people want something for nothing: Public opinion on taxes and government spending." National Tax Journal 22(supplement):113–129.
Converse, Jean (1976)
"Predicting no opinion in the polls." Public Opinion Quarterly 40:515–530.
Ferber, Robert (1966)
"Item nonresponse in a consumer survey." Public Opinion Quarterly 30:399–415.
Francis, Joe, and Lawrence Busch (1975)
"What we don't know about 'I don't knows.' " Public Opinion Quarterly 39:207–218.
Hanushek, Eric A., and John A. Jackson (1977)
Statistical Methods for Social Scientists. New York: Academic Press.
Johnson, Charles A., Roger C. Shaefer, and R. Neal McKnight (1978)
"The salience of judicial candidates and elections." Social Science Quarterly 59:371–378.
Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark Westlye, and Raymond E. Wolfinger (1977)
"The myth of the independent voter." Paper presented at the annual meeting of the American Political Science Association.
Lorge, Irving (1936)
"Prestige, suggestion, and attitudes." Public Opinion Quarterly 7:386–402.
Lovrich, Nicholas P., and Charles H. Sheldon (1983)
"Voters in contested nonpartisan judicial elections: A responsible electorate or a problematic public?" Western Political Quarterly 36:241–256.
Lovrich, Nicholas P., Charles H. Sheldon, and Erik Wassman (1988)
"The racial factor in nonpartisan judicial elections: A research note." Western Political Quarterly 41:807–816.
Pomper, Gerald (1966)
"Ethnic and group voting in nonpartisan municipal elections." Public Opinion Quarterly 30:79–97.
Public Opinion (1986)
"Opinion roundup." 9(2):21–40.
Rosen, Corey M. (1973)
"A test of presidential leadership of public opinion: The split-ballot technique." Polity 6:282–290.
Rugg, Donald, and Hadley Cantril (1944)
"The wording of questions." Pp. 23–50 in Hadley Cantril (ed.), Gauging Public Opinion. Princeton: Princeton University Press.

Schuman, Howard, and Stanley Presser (1981)
Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context. New York: Academic Press.
Sears, David O., and Jack Citrin (1982)
Tax Revolt: Something for Nothing in California. Cambridge, MA: Harvard University Press.
Sigelman, Lee, and Dan Thomas (1984)
"Opinion leadership and the crystallization of nonattitudes: Some experimental results." Polity 18:484–493.
Squire, Peverill, and Eric R. A. N. Smith (1988)
"The effect of partisan information on voters in nonpartisan elections." Journal of Politics 50:169–179.
Stagner, R. (1941)
"A comparison of the Gallup and Fortune polls regarding American intervention policy." Sociometry 3:239–258.
Stolz, Preble (1981)
Judging the judges: The investigation of Rose Bird and the California Supreme Court. New York: Free Press.
Sudman, Seymour, and Norman M. Bradburn (1982)
Asking Questions. San Francisco: Jossey-Bass.
Tarantino, Santo J., and Robert E. Jednak (1972)
"A re-examination of factors influencing the evaluation of assertions." Public Opinion Quarterly 36:109–113.
Wolfinger, Raymond E., and Steven J. Rosenstone (1980)
Who Votes? New Haven: Yale University Press.

THE POLLS—A REPORT

PUBLIC OPINION TRENDS: PRIVACY AND INFORMATION TECHNOLOGY

JAMES E. KATZ AND
ANNETTE R. TASSONE

While there have been numerous cross-sectional analyses of public opinion toward privacy (Harris and Westin, 1979), and especially computers and privacy (SNET, 1984; Gandy, 1989; Anderson, 1972), there have been few trend analyses. Consequently, we have had no good answers for the fundamental question, Is public concern over privacy rising? One answer to this question is a clear "No." Basing their position on data collected through 1983, Dutton and Meadow (1987: 168), in their definitive review of surveys on privacy, concluded that "the perceived likelihood of privacy invasions and their impact on American life has remained stable since 1974. Likewise, while computing was increasingly seen as a threat to privacy between 1973 and 1979, this trend has leveled off since then" (see also Dutton and Meadow, 1985). Yet polls conducted since Dutton and Meadow's review suggest that public concern over privacy is rising in the mid to late 1980s and what was a definitive statement a few years ago may now need to be modified. In this article we present some trends from recent poll results about privacy in general and in regard to information technology (specifically computers and telephones).

Results of Surveys

IMPORTANCE OF PRIVACY

Surveys in 1988 and 1989 show that people overwhelmingly say that privacy as an abstract concept is important (Table 1).[1] When inter-

---

JAMES E. KATZ is a sociologist at Bell Communication Research (Bellcore). ANNETTE TASSONE is completing her doctorate in human factors at Stevens Institute of Technology. The authors thank Tom W. Smith, Diane Duffy, Richard Clayton, Ed Pinnes, Bob Kraut, and Oscar H. Gandy, Jr., for their help. Maritz survey data appear in Gandy, 1990.

[1] Unlike many other topics of research, surveys of privacy concerns are by their nature likely not to include members of the population who would be most concerned about

Public Opinion Quarterly Volume 54:125–143 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-01/$2.50

viewees were asked to compare the importance of personal privacy with "other subjects on your mind" (unspecified), two in three rated it as very important, and another one in four said that it is somewhat important. Less than one in ten rated privacy as slightly or not important at all. This rating seems stable given that the results are from samples drawn in sequential years; the "very important" rating dropped 3%, but the rest of the pattern changed no more than 2%.

While people think that privacy is important as an abstract principle, it does not necessarily follow that they feel it is being threatened or invaded. That is, something can be highly valued, like the secret ballot, but citizens might not feel that the right is being eroded. Nevertheless, there is evidence showing that, since 1980, privacy invasion is of rising concern to the public (Tables 2 and 3). Two different sets of questions covering 1978 to 1983 and 1988 to 1989 both show increases in people's expressed concern over threats to, or invasions of, their privacy. Table 2 shows a change from 31% to 47% "very concerned" about their privacy between 1978 and 1983; "not concerned at all" dropped from 19% to 8% over the same time. Table 6 indicates a 6% growth in "very concerned" between 1988 and 1989.

CONCERN ABOUT THE FUTURE

Do Americans think there will be less privacy in the future? Interestingly, they say yes (compare Tables 3 and 7 for 1988). Has there been a change in this perception? Not much. When asked about future "threats" or "concerns" relative to personal privacy, one data set indicates that concern has for the most part risen since 1980, although not dramatically; the rise is no more than 6% (Table 4). Even this slight rise among those who think privacy will be a serious problem is not paralleled by a decrease in those who do not think privacy will be a problem. Tables 5 and 6, which cover the mid 1970s to the mid 1980s, show very little net rise. Although the data were not gathered annually, all three series (Tables 4, 5, and 6) suggest that pessimism over future privacy climbed during the 1970s, dipped slightly in the early eighties, and climbed again in the mid 1980s. But again the changes are undramatic.

People are more likely to say that privacy loss is going to be a "problem" (Table 4) rather than a "serious threat" (Table 5) to society. They

---

invasions of privacy. This is because potential respondents presumably would not be willing to reveal information about themselves to pollsters, judging them to be privacy invaders. Consequently, public opinion polls on the subject are probably underestimating the level of public concern. There is little practical way to surmount this weakness since standard practices of estimating nonrespondent characteristics would not work well for this dimension of nonresponse.

seem twice as likely to rate privacy as a future "problem" than as a future "threat." The relative differences lead us to conclude that privacy, while an important right, is probably a "second order" concern, coming after such problems as unemployment, crime, and nuclear war. (This hypothesis finds support in a Canadian study of relative privacy concerns, Vidmar and Flaherty, 1985, which concluded that while privacy is one of the most highly prized rights, it is a subsidiary concern to physical threats of nuclear war and street crimes.)

A major source of variance in the responses to the importance of privacy seems to be whether or not the issue of "invasion of privacy" is raised in absolute or relative terms. When it is the only issue presented, the percentage of people responding that privacy is a concern is greater than when there are many competing issues to evaluate alongside it.[2]

COMMERCIAL LIFE AND PRIVACY INVASIONS

Two samples of whether people have felt themselves to be victims of privacy invasion show little change over time; about one in five say that they have been (Table 8). When seen against the total adult population, this means that perhaps 30 million Americans have felt themselves to be so victimized. Between 1974 and 1983, there has been a large increase in the percentage of people who think they surrender their privacy when they first enter credit society (Table 9). Between 1974 and 1977 there has also been a large increase in the percent who believe organizations ask too much personal information (Table 10).

These data suggest that although people are not increasingly experiencing privacy invasions, they are increasingly feeling that consumer society is requiring a loss of privacy. Only infrequently, though, do they forgo some service to protect their privacy (Table 11). Fourteen percent of a 1978 sample said they have not applied for something like a job, credit, or insurance because they did not want to provide certain information about themselves. And the percentage of people who are willing to forgo benefits due to privacy concerns apparently has decreased since then; 11% in 1989 said they had refused to apply. Contrasting Table 11b with Table 11a shows that while people may still apply for beneficial goods like jobs and credit in large number, many also refuse to give all the information that is asked when doing so violates perceived privacy rights. From this perspective it may be said that people are devising their own resistance strategies when confronted with perceived infringements on their privacy.

---

[2] Related to this phenomenon is Dutton and Meadow's position that surveys of privacy which focus on privacy alone exaggerate respondents' levels of concerns. (Our inspection of the surveys lent credence to this prospect.) While this might have occurred, none of the Cambridge Polls, which generally show higher levels of privacy concern, focused exclusively on privacy. Indeed, Cambridge polls are house-to-house, which means they get people who do not have phones (or refuse to answer phone surveys), in contrast to most other surveys reported here, which have been done via telephone. This suggests that people who are interviewed face-to-face evince stronger privacy concerns than those who are interviewed via telephone.

We posit that people are becoming more acclimated to the necessity of giving up personal information, although they are not pleased with the prospect. Thus there may be a widening split between a sense of privacy loss, which seems to be growing, and reports of actual privacy invasion, which are stable. This schism may be attributable to generalized perceptions of social change in the former case and specific incidents in the latter case.

Vital to understanding this schism is the distinction between loss and invasion. People may choose not to apply for something because they will lose their privacy, but they may not feel that the process of applying is an invasion (since they are free not to apply). In parallel circumstances, people may feel they lose their privacy if they speak out at a public meeting, since others will thereafter know their views, but they would not consider the public meeting itself an invasion. So privacy invasions, which are seen as active events or conditions, can give rise to feelings of privacy loss, but feelings of privacy loss can be passive and not the result of a particular agent in the individual's environment taking action against the individual.

INVOLUNTARY SUBJECT OF FILES

The percentage of people who feel that they are subjects of a file for purposes unknown to themselves has risen substantially since the early 1970s (Table 12). At that time less than half the sample thought they were subjects of files (44%), but this has risen to more than two-thirds, a finding consistent since 1983. (In 1989, 68% thought they were subjects of such files.) About one in four people said they felt threatened by these files in 1974, while one in three said so in 1977 (Table 13). This increase in perceptions of being the subject of a file (Table 12) is probably correct because there has been growing computerized data collection on individuals. Justifications of threat feelings (Table 13) are, of course, more difficult to validate.

"BIG BROTHER" FEARS RISING

A noteworthy finding is the dramatic upswing in the percentage of people who state that "we are already there" when asked how close we are to a "Big Brother" style society like that described in George

Orwell's *1984* (Table 14). In the 1979–83 period, less than 10% said we were already in a society where the "government" knew "almost everything" about everyone and that virtually all personal privacy had been lost. By 1989 nearly 20% said our society had become like that. This perhaps reflects a generalized perception of privacy loss that society as a whole has experienced, rather than any personal biographical incident (as would be suggested by the rate of perceived privacy violation shown in Table 8). It may be that the revelations of the mid 1970s (Rockefeller Commission, Church hearings, FBI excesses) and, in the late 1980s, media reports on privacy loss[3] make people feel more vulnerable without necessarily having had an actual experience in this regard.

DESIRE FOR GOVERNMENTAL ACTION

The proportion of people favoring major governmental intervention to protect privacy has fallen from 54% in 1974 to 39% in 1986 (Table 15). This may be a reflection of the public's disenchantment with governmental solutions or a belief that such a major effort is not worthwhile either in general or relative to other problems. However, the mood might have shifted since 1986. Cambridge polls taken in the late 1980s reveal that about half the sample favors "new laws" (unspecified) to protect personal privacy (Table 16). (We cannot tell whether the upswing is related to a general belief that more laws are necessary or that governmental action to protect privacy itself is now viewed more favorably.) Another set of Cambridge polls, also taken in the late 1980s, shows a sharp upswing in support for a specific legislative proposal aimed at protecting privacy, i.e., to restrict exchange of information between the governmental and private sectors (Table 17). These surveys indicate that the downward trend in public support for governmental privacy laws may have been reversed and is now moving upward.

CONCERN OVER COMPUTERS

While Dutton and Meadow found that concern over privacy and computers had risen in the 1970s and subsided by the early 1980s, our evidence shows that concern might again be rising. Two longitudinal studies, 1974–83 and 1983–89, show an upward trend (Tables 18 and 19). In Table 19, concerns over safeguarding computer-stored information have eased slightly between 1988 and 1989, but they are still well above the starting point of the series, 19% in 1983.

---

[3] Our review of mass media literature shows renewed coverage of privacy invasion since the mid 1980s.

A third longitudinal series (Table 20) suggests that computer-stored information is seen as somewhat less of a threat in 1989 relative to 1986 (22% vs. 28%). Although the trend of those responding that computer-stored information represents a "very serious" threat has decreased somewhat from its high, it is still higher than when the series began in 1983. Observe that those who say "not at all" have decreased from 24% in 1983 to 17% in 1989. Note too that Table 20, which refers to "threat," is somewhat curvilinear, and Table 19, which refers to "concerns," rose steadily to a high in 1988.

Table 21 is interesting in that it shows when asked to say whether business or government poses a greater threat of abusing computers, the public increasingly volunteers the opinion "both," a choice not offered. But the public does not show a marked opinion about which sector is the larger source; both government and private sectors receive slightly more than 20% of the responses (with "government" slightly preferred as the source).

Table 22, containing data from the mid 1970s, shows a slight rise in feelings that computers have infringed on one's privacy.

TELEPHONES AND PRIVACY

Perhaps surprisingly, the public judges telephone communications as an area where privacy interests should be strongly protected. This appears to be true both for records of calls made and the content of the call itself (which could be subject to wiretapping).

Table 23 presents data from three Harris surveys (1974, 1976, and 1977) asking about the dangers of collecting personal information on computers. Of the 13 areas read to the respondents, the one that was consistently the most disapproved of was "keeping computerized records of all telephone calls made from a number," where a mean of 67% agreed with this sentiment. Keeping computerized records of telephone calls was considered even more dangerous than keeping records of political affiliations and associations, results of psychological tests, or results of intelligence tests. This table provides strong evidence that telephone privacy rights are important to the public, and that they consider information concerning their telephonic activities to be highly sensitive.

The public has been queried about wiretapping under various conditions. Since 1974 the National Opinion Research Center has asked the public about its approval of wiretapping, "everything considered" (Table 24). These data show that the public had overwhelmingly disapproved of wiretapping in 1974, to the extent of 83%. By 1977, there had been a 5% drop to 78% disapproval; since then there has been some intermittent downward fluctuation of opinion, but in 1988 it was back up at 78% again. Preliminary data for 1989 reveal a striking 9% drop in disapproval to 69%. Not surprisingly, this downward trend in disapproval is accompanied by an upward trend in approval, reaching a high of 26% in 1989. It may be that concern over drug trafficking has made people more approving of wiretapping.

The data are noteworthy in two other respects. First, despite the growing approval of wiretapping, Americans oppose it by better than two to one. We believe this lends support to the contention that telephone privacy is a highly regarded right and expectation.

Second, the upward trend in approval has been remarkably stable in its linearity, with only minuscule aberration until 1989. This would suggest that the public has relatively consistent opinions concerning wiretapping which are only changing gradually. From the vantage point of technology, such stability may be due in part to the fact that wiretapping has long been a familiar possibility. (Wiretapping was the subject of a novel as early as 1906 [Stringer, 1906].) If we were to find a dramatic change in perceptions about threats to privacy, we could anticipate that it would be associated with new and seemingly exotic technologies like computers, and such appears to be the case. Yet public understanding of the privacy implications of computer technology is inchoate, which may explain the fluctuating opinions concerning it. In marked contrast, wiretapping has widely recognizable and comprehensible connotations and opinion changes concerning it seem not at all erratic.

The extended run of data on wiretapping in Table 24 probes general approval of the practice. When wiretapping is presented to respondents in particular contexts, which is done in Tables 25–28, the answers vary somewhat but still show marked disapproval. This we believe constitutes additional support for our contention that telephone privacy rights are highly regarded by the American public. For example, when asked about wiretapping as a way to combat crime, the public in a 1981 survey still disapproved of it by a two-to-one ratio (Table 25).

The characteristics of the target of a proposed wiretap affect approval ratings. A 1986 survey showed that when wiretapping is used (without a court order) to gather evidence against a suspect with a criminal record versus a suspect without one, the low rate of approval of such surveillance more than doubles in the former case ("definitely" allowing wiretap increased to 22% from 8% if the suspect had a criminal record; Table 26). Notably, half the respondents believe wiretapping of a suspect without a court order should not be allowed even if the suspect has a long criminal record; three quarters are against it in the absence of a long criminal record. This can be adduced as persuasive evidence of the public's disapproval of wiretapping.

Several survey questions combined wiretapping and electronic surveillance and asked the public to judge their permissibility in criminal investigations. Combining them makes it more difficult to assess what the public thinks, since electronic surveillance means a wide variety of things, from attaching monitors under car bumpers to "bugging" a bedroom. Nonetheless, judgments can be made. For example, due process appears to be an important concept in the public's mind. Respondents were asked in 1974 and 1976 about a scenario in which the government would be allowed to wiretap and surveil either with or without court permission. There is much higher approval with court review than without court review (Tables 27a and 27b). Certainly the specific nature of the question's wording and the hypothesized situation affects approval rates. But significantly, these data reveal the high value of due process to Americans when telephone privacy invasions are at risk.

The public's basic disapproval of wiretapping and surveillance is reflected in Table 28. In 1974 and 1976 about half the sample disapproved of the government *ever* collecting data by these means. The confounding effect of including wiretapping and electronic surveillance in the same question may be revealed by contrasting Table 28 with Table 24. One possible way to explain the difference in approval rates of surveillance (~48% vs. ~18%) is that Table 28 includes "electronic surveillance" in its question. (Other wording differences, such as the word "ever" in Table 28, as well as survey method, could also account for the difference. It seems unlikely that the shift could be attributable to attitude changes in the population.) Contrasting the questions in these tables underscores the way in which attributes of the invasion target and procedures used in the proposed invasion significantly affect the public's evaluation of their permissibility.

Table 29 changes the focus from approval of wiretapping to perceived victimization rates by presenting the results of two polls taken in January and December of 1978. It shows that, at the end of 1978, 9% believed their phones had ever been tapped, a 3% decrease from the beginning of 1978, when 12% said they believed this had occurred. There was virtually no change in the 85% response level among those who did not think their phone had ever been tapped. The result that about 10% of the population thinks they have had their phone tapped implies a widespread belief of being subjected to tapping (and by extension that wiretapping is a widespread activity); however by no means does the bulk of the population hold this view.

Summary and Conclusion

Longitudinal public opinion surveys provide evidence that

1. Privacy, as an abstract concept, is one that most Americans value and judge important, although we found only slight evidence that people will sacrifice to protect it
2. There is modest and often conflicting evidence that there has been an upswing in privacy concern in the 1980s
3. Americans believe privacy loss will be a larger problem in the future than it is today
4. The proportion of people who believe they give up their privacy to participate in consumer society has grown, as has the proportion who believe they are the subject of computer files being put together for purposes unknown to them
5. Although people are increasingly willing to try to get tangible benefits even if it requires divulging privacy-sensitive information, a substantial percentage are refusing to supply complete information on themselves
6. Public support for major governmental actions to protect privacy have decreased, but recently there is a rise in support for some particular laws to protect privacy
7. The way in which an issue is presented consistently influences the response (e.g., "concern" vs. "threat" to privacy)
8. Telephone privacy is highly regarded and should be protected by due process
9. There is consistency in the strong disapproval of wiretapping, although people increasingly approve of the practice

Even though Americans perceive that their privacy is being threatened, little evidence was found that most Americans forgo major benefits even if getting them violates their perceived privacy; they do, however, appear to try to get the benefits while not disclosing all the information that is asked in an application and thus act to some extent in accordance with their beliefs. For example, while they believe that credit card companies require privacy loss, relatively few are apparently willing to give up such conveniences in order to protect their personal privacy. Most of the data to support this conclusion were collected in the late 1970s; yet with the spread of credit card usage, it is likely that this description is even more apropos. Still, while people may be somewhat more likely to apply for jobs, credit, or insurance even if doing so requires divulging personal information, great numbers of them also appear to withhold some privacy-sensitive information asked for in such applications. Again, it is not clear whether people may be actually feeling privacy "loss" when they agree to the sentiment that they have experienced "invasion" at the hands of a credit card agency.

About 18% of Americans report that they have been victims of privacy invasion. Apparently privacy invasion is not something most Americans experience in their daily lives, despite their seeing it as a real concern (albeit often an abstract one). It bears restating that our data address perceptions and attitudes almost exclusively, rather than reports of actual behavior. In this regard there certainly may be severe privacy violations occurring that few are aware of or define as such.

Another noteworthy finding can be seen in four tables: the percentage of uncertainty in Americans' responses, e.g., "not sure" or "don't know," has decreased for many types of questions (Tables 4, 9, 18, 19). There was no systematic increase in "don't knows" for any of the questions. This suggests that some Americans have been making up their minds about privacy, and when they do so, it is toward greater concern.

Dutton and Meadow (1986) found that privacy concern in some areas increased after Watergate in the mid 1970s and then subsided. It is noteworthy, though, that there are several pieces of evidence that point to concerns rising again as America has entered the late 1980s, particularly in the area of computers. The most recent wave of concern might be related to filtering down of computers to the mass level, where, according to a 1989 Cambridge Reports survey, today nearly 30% of the population uses computers directly in their work or education and thus are familiar to an extent with their power and purposes. These pieces of evidence support a "wave" hypothesis, which holds that concerns rise, then fall, intermittently in response to headlines as well as political and technological change. Yet few other findings support the wave hypothesis, perhaps because of inconsistency in question wording from year to year. Overall, the evidence supporting a wave hypothesis is meager. This leads us to think that any incident "setting the privacy wave in motion" would have to be quite profound—probably of the magnitude of Watergate—in order for it to be observable as an isolated cause of opinion change.

The growth in privacy concern since the mid 1980s, given the major changes in the power and distribution of information processing technology, is not surprising. Yet attributing a public attitude change to one factor alone is not our position, and we recognize the difficulty of separating opinion changes due to computerization from other societal factors. Of continuing interest are (1) the direction of change (if any) in this level of concern as technology continues to advance and (2) if the direction of change is up, the manner in which public opinion will affect policy and social practice.

Cambridge—Cambridge Reports, Inc., Cambridge, MA
Gallup—The Gallup Organization

Harris—University of North Carolina, Chapel Hill (Social Science Data Library)
Maritz—Maritz Marketing Research, Inc., St. Louis
NORC—National Opinion Research Center, University of Chicago
Roper—Roper Center for Public Opinion Research, University of Connecticut, Storrs

References

Anderson, Ronald E. (1972)
"Sociological analysis of public attitudes toward computers and information files." Pp. 649–675 in Proceedings of the Joint Computer Conference. Montvale, NJ: American Federation of Information Processing Societies.

Dutton, William H., and Robert G. Meadow (1985)
"Public perspectives on government information technology." Unpublished report for the Office of Technology Assessment, U.S. Congress.

——— (1987)
"A tolerance for surveillance: American public opinion concerning privacy and civil liberties." Pp. 147–170 in Karen B. Levitan (ed.), Government Infostructures. New York: Greenwood Press.

Gandy, Oscar (1989)
"Information privacy and the crisis of control." Pp. 59–73 in Marc Raboy and Peter Bruck (eds.), Communications: For and Against Democracy. Montreal: Black Rose Books.

——— (1990)
"Caller identification: The two-edged sword." Paper presented at the Pacific Telecommunication Council 12th annual conference, Honolulu, HI, 15 January.

Harris, Louis, and Alan Westin (1979)
The Dimensions of Privacy. Stevens Point, WI: Sentry Insurance.

Southern New England Telephone (SNET) (1984)
The Road After 1984. Poll conducted by Louis Harris. New Haven, CT.

Stringer, Arthur (1906)
The Wire Tappers. Boston: Little, Brown.

Vidmar, Neil, and David Flaherty (1985)
"Concern for personal privacy in an electronic age." Journal of Communication 35:91–103.

1. *Cambridge:* Compared with other subjects on your mind, how important is personal privacy: Is it very important, somewhat important, not too important, or not important at all?

	Q1/88	*Q1/89*
Very important	70%	67%
Somewhat important	23	24
Only slightly important	4	4
Not important at all	2	3
Not sure	1	3
	100	101
N	1,471	1,448

2. *Harris:* How concerned are you about threats to your personal privacy in America today? Would you say you are very concerned, somewhat concerned, only a little concerned or not concerned at all?

	11/78	*7/82*	*9/83*
Very concerned	31%	45%	47%
Somewhat concerned	33	29	30
Only a little concerned	17	14	15
Not concerned at all	19	11	8
DK (vol.)	1	1	0
	101	100	100
N	1,256	1,513	1,506

3. *Cambridge:* How concerned are you about the invasion of your personal privacy in the United States today? Are you very concerned, somewhat concerned, only a little concerned or not concerned at all?

	Q1/88	*Q1/89*
Very concerned	38%	44%
Somewhat concerned	37	32
Only a little concerned	18	17
Not concerned at all	7	6
DK (vol.)	1	1
	101	100
N	1,471	1,448

4. *Roper:* Here is a list of some different kinds of problems people might or might not be facing 25 to 30 years from now. Would you go down that list, and for each one tell me whether you think it will or will not be a serious problem your children or grandchildren will be facing 25 to 50 years from now. . . . Lack of privacy (government surveillance, data files, etc.)

	1/75	*1/77*	*1/79*	*2/81*	*2/83*	*2/85*
Will be serious problem	55%	56%	59%	55%	63%	61%
Will not be serious problem	30	30	30	35	27	32
Don't know	15	14	11	11	10	7
	100	100	100	101	100	100
N	2,005	2,000	1,997	2,000	2,000	1,977

5. *Roper:* Now here is a list of a number of different things that people said may be happening in our society in the coming years. Would you read down that list and call off all those you think are likely to happen in the coming years?

	5/74	*5/76*	*5/78*	*4/80*	*5/84*
Increasing invasion of privacy	50%	53%	54%	50%	54%
N	1,984	2,002	2,003	2,002	2,000

6. *Roper:* Even though some of those things are likely to happen, you may feel they will have different effects on our society. Would you read down that list again and tell me which ones you see as serious threats to our society and life as we know it in the United States.

	5/74	*5/76*	*5/78*	*4/80*	*5/84*
Increasing invasion of privacy	25%	30%	30%	22%	27%
N	1,984	2,002	2,003	2,002	2,000

7. *Cambridge:* What about in future years—how concerned are you about the invasion of your personal privacy in the future: are you very concerned, somewhat concerned, not too concerned, or not concerned at all?

	Q1/88
Very concerned	45%
Somewhat concerned	35
Only a little concerned	13
Not concerned at all	6
DK (vol.)	1
	100
N	1,513

8. *Harris:* Have you personally ever been the victim of what you felt was an improper invasion of privacy?

	11/78	*9/83*
Yes, have been	19%	20%
No, have not been	78	78
Not sure	3	2
	100	100
N	1,513	1,256

9. *Harris:* Some people say that Americans begin surrendering their privacy the day they open their first charge account, take out a loan, put something on the installment plan, or apply for a credit card. All in all, do you tend to agree or disagree with this statement?

	3/74	*1/76*	*3/77*	*6/78*	*11/78*	*9/83*
Agree	48%	47%	67%	71%	76%	69%
Disagree	43	47	25	24	21	30
Not sure	9	6	8	5	3	1
	100	100	100	100	100	100
N	1,495	1,532	1,522	1,259	1,513	1,256

10. *Harris:* Do you feel that organizations and agencies ask you for too much personal information, or don't you feel that way?

	3/74	*6/76*	*3/77*
Feel that way	34%	33%	59%
Don't feel that way	59	60	34
Not sure	7	7	7
	100	100	100
N	1,495	1,532	1,522

11a. *Harris, Maritz:* Have you ever decided not to apply for something like a job, credit, or insurance, because you did not want to provide certain kinds of information about yourself?

	Harris *11/78*	*Maritz* *1–2/89*[a]
Yes, did not apply	14%	11%
No	85	89
Can't remember, not sure, DK	1	0
	100	100
N	1,513	981

[a] Have you ever refused to apply for something like a job, credit, or insurance because you did not want to provide all the information you thought they would ask?

11b. *Maritz:* Have you ever refused to provide information on an application that you felt they had no right to ask?

	1–2/89
Yes	35%
No	67
Don't know	0
	100
N	981

12. *Harris:* Do you believe that personal information about yourself is being kept in some files somewhere for purposes not known to you, or don't you believe this is so?

	3/74	*1/76*	*1/77*	*9/83*	*Q1/88*[a]	*Q1/89*[a]
Believe	44%	47%	54%	68%	69%	68%
Do not believe	44	43	32	30	17	20
Not sure	12	9	14	2	13	12
	100	99	100	100	99	100
N	1,495	1,532	1,522	1,256	1,471	1,448

[a] *Cambridge:* . . . not known to you or not?

13. *Harris:* Do you feel threatened in any way by having information about yourself in some files or don't you feel threatened by that?

	3/74	*1/76*	*1/77*
Feel threatened	24%	26%	32%
Not threatened	73	70	62
Not sure	3	4	6
	100	100	100
N	1,495	1,576	1,522

14. *Harris:* In a book, *1984*, by George Orwell, virtually all personal privacy had been lost and the government—called "Big Brother"—knew almost everything that everyone was doing. Whether or not you have read the book, how close do you think we are to that kind of a society—are we there already, very close, somewhat close or not close at all?

	1979	*1983*	*Q1/88*[a]	*Q1/89*[a]
Already there	8%	6%	16%	19%
Very close	26	23	28	23
Somewhat close	39	40	34	31
Not close at all	19	28	15	19
Not sure (don't know)	8	3	8	7
	100	100	101	99
N	1,513	1,256	1,471	1,448

[a] *Cambridge:* Some people are worried about our society becoming one in which individuals have lost almost all their privacy and the government knows almost everything everyone is doing. How close do you think we are to that kind of society: Are we already there, very close, somewhat close, or not close at all?

15. *Roper:* There are many problems facing our nation today, all of which the federal government must be concerned with. But at certain times some things

are more important than others, and need more attention from our government than others. I'd like to know for each of the things on this list whether you think it is something the government should be making a major effort on now, or something the government should be making some effort on now, or something not needing any particular government effort now. . . . Seeking ways to protect the privacy of individuals in our society

	1974	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*
Major effort	54%	56%	56%	51%	44%	46%	43%
Some effort	31	32	31	34	36	37	38
Needs no particular effort now	11	8	9	12	15	13	15
Don't know	3	4	4	4	5	4	4
	99	100	100	101	100	100	100
N	1,987	2,004	2,004	2,001	2,002	2,006	2,006

	1981	*1982*	*1983*	*1984*	*1985*	*1986*
Major effort	44%	38%	43%	46%	48%	39%
Some effort	35	39	36	35	37	43
Needs no particular effort now	16	19	16	15	13	13
Don't know	4	4	4	4	3	5
	99	100	99	100	101	100
N	2,003	2,000	2,000	2,000	2,000	1,994

16. *Cambridge:* Do you think we need new laws to protect personal privacy, or are existing laws adequate?

	Q1/88	*Q1/89*
Need new laws	48%	50%
Existing laws adequate	37	37
Don't know	15	13
	100	100
N	1,471	1,448

17. *Cambridge:* Do you favor or oppose laws that would restrict the exchange of information between government and private institutions?

	Q1/86	*1987*	*Q1/88*
Favor	56%	—	69%
Oppose	32	—	23
DK (vol.)	12	—	9
	100		101
N	1,471		1,448

18. *Harris:* In your opinion, do you feel that the present uses of computers are an actual threat to personal privacy in this country, or not?

	1974	*1976*	*1977*	*1978*	*1983*
Are a threat	38%	37%	41%	54%	52%
Not a threat	42	51	44	32	43
Not sure	20	12	15	14	5
	100	100	100	100	100
N	1,495	1,532	1,522	1,513	1,256

19. *Cambridge:* How concerned are you about safeguarding the privacy of information about you that is stored in computers? Are you very concerned, somewhat concerned, not too concerned, or not concerned at all?

	1983	*1984*	*1985*	*1986*	*1987*	*1988*	*1989*
Very concerned	19%	26%	26%	28%	—	41%	35%
Somewhat concerned	27	29	28	31	—	34	38
Not too concerned	24	18	16	18	—	15	14
Not concerned at all	18	17	14	18	—	8	11
Don't know (vol.)	12	10	16	10	—	2	2
	100	100	100	105		100	100
N	1,466	1,429	1,430	1,455		1,471	1,448

20. *Cambridge:* As computer usage increases in business and the general society, more and more information on individual consumers is being acquired and stored in various computers. How serious a threat to your personal privacy is this development? Is it a very serious threat, somewhat serious, only slightly serious, or not a serious threat to your privacy at all?

	1983	*1984*	*1985*	*1986*	*1988*	*1989*
Very serious	19%	26%	26%	28%	23%	22%
Somewhat serious	27	29	28	31	38	33
Slightly serious	24	18	16	18	19	23
Not at all	24	21	22	18	14	17
Don't know	5	5	8	5	6	5
	99	99	100	100	100	100
N	1,466	1,429	1,430	1,455	1,471	1,448

21. *Cambridge:* To the extent that the increased use of computers does pose a threat to personal privacy, do you think the greater threat comes from the improper use of computers by businesses or by government?

	1985	*1986*	*Q1/88*	*Q1/89*
Business	22%	22%	21%	20%
Government	26	26	24	22
Both (vol.)	35	40	42	44
Neither (vol.)	7	6	5	6
Other (vol.)	2	—	1	1
Don't know (vol.)	9	6	8	7
	101	100	101	100
N	1,430	1,455	1,471	1,448

22. *Harris:* Have you ever felt computers that store a lot of information about you infringed on your right to privacy or not?

	1974	*1976*
Have felt	35%	39%
Have not	57	54
Not sure	8	7
	100	100
N	1,495	1,532

23. *Harris:* Now I would like to ask you about some types of information that have been suggested for collection and storage in computers. For each information area I call off, tell me if you feel that the benefits outweigh the dangers in keeping that type of information in a computer, or that dangers outweigh the benefits. . . .

	1974	*1976*	*1977*
A student's academic record			
Benefits outweigh dangers	70%	69%	66%
Dangers outweigh benefits	18	21	23
Not sure	12	10	11
A worker's employment record			
Benefits outweigh dangers	68%	69%	64%
Dangers outweigh benefits	22	23	26
Not sure	10	8	10
Intelligence test scores			
Benefits outweigh dangers	54%	51%	50%
Dangers outweigh benefits	35	39	37
Not sure	12	10	13
Results of psychological tests			
Benefits outweigh dangers	45%	41%	42%
Dangers outweigh benefits	41	45	43
Not sure	14	13	15
Police record of any person arrested on suspicion of a crime			
Benefits outweigh dangers	—	38%	41%
Dangers outweigh benefits	—	55	51
Not sure	—	7	8
Police records of any person arrested and convicted			
Benefits outweigh dangers	71%	84%	83%
Dangers outweigh benefits	22	10	11
Not sure	7	6	6
The political affiliations and associations of a person			
Benefits outweigh dangers	29%	24%	25%
Dangers outweigh benefits	59	66	62
Not sure	12	10	13
An individual's medical record			
Benefits outweigh dangers	75%	66%	69%
Dangers outweigh benefits	18	28	23
Not sure	7	6	8
A complete history of a person's traffic violations			
Benefits outweigh dangers	56%	50%	49%
Dangers outweigh benefits	34	42	40
Not sure	10	8	11
A complete record of all the telephone calls made from a number			
Benefits outweigh dangers	24%	21%	26%
Dangers outweigh benefits	64	72	61
Not sure	12	7	13
Complete credit information about a person			
Benefits outweigh dangers	55%	53%	52%
Dangers outweigh benefits	36	37	37
Not sure	9	9	11
Weapons owned by an individual			
Benefits outweigh dangers	63%	55%	51%
Dangers outweigh benefits	29	38	36
Not sure	8	7	13
Mental health record of an individual			
Benefits outweigh dangers	51%	46%	48%
Dangers outweigh benefits	39	44	40
Not sure	10	9	12
N	1,495	1,532	1,522

24. *NORC:* Everything considered, would you say that in general you approve or disapprove of wiretapping?

	1974	*1975*	*1977*	*1978*	*1982*	*1983*	*1985*	*1986*	*1988*	*1989*
Approve	17%	16%	18%	19%	19%	19%	23%	22%	22%	26%
Disapprove	80	80	78	78	77	78	74	74	78	69
DK	4	4	3	3	4	4	3	4	1	5
	101	100	99	100	100	101	100	100	101	100
N	1,484	1,490	1,530	1,532	1,506	1,599	1,534	1,470	927	1,000

25. *Gallup:* Do you approve or disapprove of these alternatives for dealing with crime. . . . Allowing the police to wiretap telephones of anyone they suspect?

	1981
Approve	30%
Disapprove	67
Don't know	3
	100
N	1,030

26. *NORC:* Suppose the police get an anonymous tip that a man . . . (with/without) . . . a long criminal record is planning to break into a warehouse. Please circle a number for each question to show if you think police should be allowed without a court order to tap his telephone.

	1985	
	With	*Without*
Definitely allowed	22%	8%
Probably allowed	22	13
Probably not allowed	24	25
Definitely not allowed	28	50
Can't choose	4	4
	100	100
N	657	660

27a. *Harris:* Do you feel the federal government should be allowed to engage in wiretapping and electronic surveillance, if in each case it had to go to court beforehand to obtain court permission or don't you feel the federal government should ever be allowed to engage in wiretapping or electronic surveillance?

	1974	*1976*
Should be allowed	63%	65%
Should not be allowed	28	27
Not sure	9	8
	100	100
N	1,495	1,532

27b. *Harris:* . . . even if a court doesn't feel it is needed?

	1974	*1976*
Should be allowed	27%	31%
Should not be allowed	68	65
Not sure	5	4
	100	100
N	1,495	1,532

28. *Harris:* Do you think the government should be able to gather evidence against citizens suspected of criminal activities by means of wiretapping and electronic surveillance, or don't you think the government should ever be able to do this?

	1974	*1976*
Should be able	41%	43%
Should not be able	47	48
Not sure	12	9
	100	100
N	1,495	1,532

29. *Harris:* Do you believe that your telephone has ever been tapped or not?

	1/78	*12/78*
Yes	12%	9%
No	86	85
Not sure/DK	2	7
	100	101
N	1,458	1,505

Public Opinion Quarterly

[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible] for Contra Aid
[illegible] F. BISHOP	Issue Involvement and Response Effects
RAYMOND A. JUSSAUME, JR. and YOSHIHARU YAMADA	Viability of Mail Surveys in Japan and the United States
STEPHEN A. AYIDIYA and MCKEE J. MCCLENDON	Response Effects in Mail Surveys
CLYDE WILCOX	Race Differences in Abortion Attitudes
THE POLLS	Review: National Crime Survey Redesign *by Wesley G. Skogan*
	Report: National Service *by Tom W. Smith*
BOOKS	Reviews and Notes *by Baruch Fischhoff, James L. Gibson, Stephen J. Cutler, and Gregory Andrade Diamond*

Summer 1990

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1987–91)
Herbert Abelson (1988–92) David L. Sills (1986–90)
Gladys Engel Lang (1989–93)

Norman Bradburn and Andrew Kohut,
AAPOR Councillors-at-Large

Howard Schuman, *Editor*

Manuscript Editor John D. Vander Weg
Office Manager Carlotta Nosel
Assistant to the Editor Cheryl Rieger

The Polls
Reviews: Stanley Presser, Seymour Sudman
Reports: Tom W. Smith

EDITORIAL BOARD

Robert P. Abelson
Barbara A. Bailar
Sandra J. Ball-Rokeach
Larry M. Bartels
George F. Bishop
Lawrence D. Bobo
Jean M. Converse
Thomas D. Cook
Irving Crespi
James A. Davis
Stanley Feldman
Richard J. Fox
Norval D. Glenn
Robert M. Groves
Deborah R. Hensler
Shanto Iyengar
Jon A. Krosnick
Kurt Lang
Michael MacKuen
McKee J. McClendon
Jack M. McLeod
John Modell
John Mueller
Harold Nisselson
Helmut Norpoth
Steven J. Rosenstone
Norbert Schwarz
David O. Sears
Robert Y. Shapiro
Lee Sigelman

Public Opinion Quarterly (ISSN 0033-362X) Volume #54, 4 issues (1990)

© 1990 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Avenue, Chicago, Illinois 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $40.00; individuals 1 year $20.00. WAPOR member subscription rate, U.S.A.: individuals 1 year $18.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $18.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

PUBLIC OPINION QUARTERLY

VOLUME 54 / SUMMER 1990 / NUMBER 2

Articles

The Threshold of Public Attention 159
W. RUSSELL NEUMAN

The Decline of Over-Time Comparability in the National Election Studies 177
PAUL R. ABRAMSON

Response: Continuity in NES Studies 191
WARREN E. MILLER

Question Wording and Public Support for Contra Aid, 1983–1986 195
BRAD LOCKERBIE and STEPHEN A. BORRELLI

Issue Involvement and Response Effects in Public Opinion Surveys 209
GEORGE F. BISHOP

A Comparison of the Viability of Mail Surveys in Japan and the United States 219
RAYMOND A. JUSSAUME, JR., and YOSHIHARU YAMADA

Response Effects in Mail Surveys 229
STEPHEN A. AYIDIYA and MCKEE J. MCCLENDON

Race Differences in Abortion Attitudes: Some Additional Evidence 248
CLYDE WILCOX

The Polls

Review: The National Crime Survey Redesign 256
WESLEY G. SKOGAN

Report: National Service 273
TOM W. SMITH

Books

Robert Cameron Mitchell and Richard T. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method* 286
BARUCH FISCHHOFF

Thomas R. Marshall, *Public Opinion and the Supreme Court* 289
JAMES L. GIBSON

M. Powell Lawton and A. Regula Herzog, eds., *Special Research Methods for Gerontology* 291
STEPHEN J. CUTLER

Book Note: Thomas W. Graham, *American Public Opinion on NATO, Extended Deterrence, and Use of Nuclear Weapons: Future Fission?* 293
GREGORY ANDRADE DIAMOND

NOTICE TO CONTRIBUTORS

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1990 issue of POQ.

QUESTION WORDING AND PUBLIC SUPPORT FOR CONTRA AID, 1983–1986

BRAD LOCKERBIE AND
STEPHEN A. BORRELLI

Abstract How does question wording affect the results of polls? How can we distinguish question wording effects from genuine change? In this analysis, we make use of a series of polls concerning President Reagan's policy toward Nicaragua. We identify five variations in wording in the survey items that have been used on this topic, and we find evidence that these are responsible for variations in support for the Contras. If the ideological nature of the Sandinistas, President Reagan, the amount of money to be spent, and/or the role of the Contras are mentioned, or if a balanced question format is employed, the poll results can be substantially affected. Once these factors are controlled for, we can argue that there was a small but statistically significant increase in support for Reagan's policy over the period 1983–1986.

One of the most vexing problems facing poll watchers is the comparison of poll results over time and across polls. An examination of two or more polls on the same general topic can show widely disparate results (Borrelli, Lockerbie, and Niemi, 1987). In the popular media, one can see that different survey organizations employing different questions come up with very different results.[1] Without taking the differences over time and across polls into account, any comparisons that are made rest upon somewhat shaky ground. In this research, we attempt to demonstrate the strong effects of question wording on aggregate public opinion regarding the funding of the Contras.

Over the past few years, the debate over U.S. government support for the Nicaraguan Contras has been one of the most tendentious in American politics. People on both sides, using different polls, have claimed that public opinion tends to support their position. How can it

BRAD LOCKERBIE is Assistant Professor of Political Science at the University of Georgia.
STEPHEN A. BORRELLI is Instructor of Political Science at the University of Alabama.

---

[1] See, for example, *U.S. News and World Report*, 11 July 1987, 44–47.

Public Opinion Quarterly Volume 54:195–208 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-02/$2.50

be that people appear, at the same time, to favor funding the Contras and also to be opposed to funding the Contras? Well, as we know, subtle variations in question wording can seriously affect the results of a poll (e.g., Mueller, 1973; Bishop, Tuchfarber, and Oldendick, 1978; Bishop, 1987). In short, the potential exists for individuals to construct survey items to favor their own side. While we do not think this is done here, it is important to demonstrate how even subtle variations in question wording can provide (unintentionally) differing results. It is also important to attempt to estimate the size of these question wording effects, and to identify the specific variations in wording that appear to bring them about.

From 1983 through 1986, the percentage favoring funding of the Contras has varied considerably, ranging from a low of 13% to a high of 42%. Appendix 1 presents a listing of the polls employed in this analysis, the percentages favoring the funding of the Contras, the percentages opposing funding of the Contras, the percentages not knowing whether the Contras should be funded, as well as the exact wording of the questions used by each poll.[2] Is this variance primarily a response to changing political evaluations or is it primarily a function of survey construction?

The questions used to ascertain public opinion on the funding of the Contras have varied considerably across polling organizations and across time, but there are enough similarities that we can identify key characteristics that might influence the results of a poll. It should be emphasized that during the period under study (1983 to mid 1986) the public's knowledge of the Nicaraguan conflict, and the United States role in it, was still in a very formative stage. A Gallup survey administered on 8–11 March 1985, for example, indicated that 48% of those sampled had followed "recent developments in Nicaragua . . . not at all closely," and that only 38% of the sample could correctly identify which side the United States was supporting. We argue that in this context of low information, the public should be especially attentive and responsive to positive or negative cues within the texts of survey questions that might indicate to them the "correct" position to take on the issue of aid to the Contras. Below, we identify five major cues whose presence or absence in our sample of question items are expected to influence public support for the Contras in a predictable way.

First, we suspect that a survey item explicitly associating President

---

[2] The data employed in the analysis are made available by the Roper Center at the University of Connecticut. No polls after mid 1986 are included, for fear that the analysis would be contaminated by the Iran-Contra scandal. In more recent polls, people may be giving their opinions on the scandal instead of responding to Contra funding. All interpretations herein are those of the authors. For the complete and exact wording of the survey items used, along with the coding decisions made by the authors, see Appendix 1.

Reagan with Contra aid would evoke a higher level of support for the program than would be expressed in the absence of his imprimatur. Several studies employing a split-ballot technique have reported that respondents react more favorably to survey items mentioning presidential endorsement of a given policy than to similar items that do not mention the president's support of the policy (e.g., Rosen, 1973; Thomas and Sigelman, 1985). For testing the hypothesis that the association of Contra funding with President Reagan increases support for Contra aid, we create a dummy variable (Reagan mention) in which all polls that mention Reagan are coded 1 and the others are coded 0.

One concern in American politics is the likelihood of the United States going to war. The Reagan administration has taken great pains to argue that funding of the Contras makes it less likely that U.S. troops will be fighting and dying in Nicaragua (Speakes, 1984; Congressional Quarterly Almanac, 1985). If the respondents are explicitly told that somebody other than Americans will be doing the actual fighting, it is likely that support for the funding will be higher than it would otherwise be. Additionally, being told that there is internal opposition to the Sandinista regime may increase the likelihood that people will support the policy, since it is less likely to appear to be a U.S. effort with no support within Nicaragua. The questions that mention the Contras (Contra mention) are coded 1, and the questions with no mention of the rebels are coded 0.[3]

A third variation on question wording that may influence support for funding the Contras is the way in which the Nicaraguan government is described. On one side, the Sandinistas are either not mentioned at all or are referred to as "the government in Nicaragua," "government troops," or "the Sandinista government in Nicaragua." In the remaining polls, the Sandinistas are identified as being "leftists," "Marxist," or "pro-Soviet." A good deal of the Reagan administration's campaign on behalf of the Contras has been devoted to emphasizing the Marxist-Leninist character of the Sandinista regime and its close relations with the Soviet Union (Kenworthy, 1987). For testing the hypothesis that identifying the Sandinistas as leftist, Marxist, or pro-Soviet increases support for the Contras, we have coded the survey items that label the Sandinista government as any of the above as 1 (Ideolog. label) and all others as 0.

All of the cues examined to this point—mention of Reagan, mention of the Contras, and ideological labeling of the Sandinistas—are expected to exert a positive impact on public support for the Contras. Two additional cues may well reduce levels of support for funding the

---

[3] The polls that mention the Contras, rebels, or forces in Nicaragua fighting the Sandinistas are all coded 1.

Contras. First, specifying the amount of money to be spent by the U.S. government should act to depress levels of support for the Contras. Those who favor the U.S. government providing some money, but not as much as the actual amount being discussed, may state that they object to funding the Contras. Questions that mention the amount of money to be spent are coded 1 (Dollar amount), and all others are coded 0.

Second, it has been shown that balanced question formats, giving each side of an issue roughly equal attention, can yield vastly different results from questions that merely ask the respondents to react negatively or positively to one side of an issue (Yeric and Todd, 1983). If the question format is balanced, opponents of Contra aid who might otherwise hide their views for fear of holding unpopular or unreasonable opinions would be encouraged to express their true opinion on the issue. Moreover, the balanced format (see Appendix 1 for actual wording) implies there is no necessary connection between opposition to the Sandinistas and support for funding of the Contras. The variable representing this (Balanced format) is coded 1 if the question employed is balanced and 0 if the question is not balanced.[4]

Of course, we do not believe that the reported variations in support for the Nicaraguan Contras are entirely the result of methodological artifacts. A cursory examination of the percentage favoring the funding of the Contras shows that there has been a modest increase over time. A time-trend variable (Time trend) that increases by one unit every day has been created; the closing date of the earliest survey in our sample is day one, and all other surveys are coded according to their closing dates. The relationship between this time-trend variable and the proportion favoring the U.S. government funding of the Contras is modest but statistically significant. The regression slope of the percentage favoring Contra aid on the time-trend variable is .009 (significant at .01, two-tailed) and the R^2 is .28. To control for this apparent shift in public opinion, we have included the time-trend variable alongside the methodological variables.

We perform multivariate ordinary least squares (OLS) regression analysis predicting a poll's reported percentage expressing support for aid to the Nicaraguan resistance from that poll's scores on the five methodological dummy variables described above and the time-trend

---

[4] For the exact wording of the "balanced" item and of all the items utilized in our sample, see Appendix 1. We should also note that we attempted to include an item representing whether or not the word "overthrow" was used in describing U.S. (and Contra) goals in Nicaragua, but we expected that it would have little additional explanatory power given the items already mentioned. As it turns out, the "overthrow" item does not attain the .20 two-tailed significance level when added to the model presented in Table 1.

Table 1. OLS Regression of Percentage Supporting Contra Aid on Question Wording Variables

Variable	Coefficient	Std. Error	*T* Ratio	Beta
Reagan mention	4.92	2.24	2.33**	.37
Contra mention	8.98	2.10	4.27***	.54
Ideolog. label	4.38	2.96	1.48*	.31
Dollar amount	−3.94	2.33	−1.69*	−.21
Balanced format	−8.58	3.87	−2.22**	−.55
Time trend	.008	.0023	3.51***	.50
Constant	13.40	2.24	5.99***	—

NOTES: R^2 = .74; adjusted R^2 = .67; N = 28; $D.W.$ = 2.21. Significance tests are two-tailed.

*$p < .20$. **$p < .05$. ***$p < .01$.

variable. To repeat, the dependent variable is the reported percentage in favor of the Reagan administration's policy with regard to funding the Contras, and the independent variables are the question wording cues: our expectation is that mention of Reagan, mention of the Contras, ideological labeling of the Sandinistas, and the time-trend variable will be positively associated with the percentage expressing support, while the balanced question format and the mention of a specific dollar amount will be negatively associated with support for the administration's position.[5]

The OLS regression results are displayed in Table 1. We find that the methodological variables, together with the time-trend variable, account for nearly three-quarters of the variance in support for Contra aid. The coefficient for the time-trend indicates that the overall increase in support for Contra aid persists even after controlling for the question-wording effects. On average, support for Contra aid increased at a rate of .008 percentage points per day (or approximately a percentage point every four months).

The performance of the question-wording items, however, leaves no doubt that methodology must be taken into account in any discussion of trends in support for Contra aid. All of the coefficients for the question-wording items are in the expected directions (positive for the first three items, negative for the last two). Three of the five coefficients for the methodology items are statistically significant at the conventional (.05) level, or better; a fourth is extremely close to conventional significance,[6] and there is less than a one in five probability that the fifth does not have the anticipated effect. Judging by the beta weights in Table 1, three of the five question-wording cues—mention of Reagan, mention of the Contras, and the balanced question format—are at least as important in accounting for support as the overall time trend.[7]

---

[5] Concern that the time-trend variable would lead to greater opposition as well as greater support, due simply to the decreasing number of don't knows, led us to run the same regression model as that presented in Table 1 with percentage opposing Contra aid as the dependent variable. When this analysis is run, we find that the level of opposition does decrease over time.

The interpretation of the coefficients is straightforward, and they help to reveal how sizable the effects of question wording are. All else being equal, mentioning Reagan's name is associated with an estimated 4.92-percentage-point increase in support for his policy. Mention of the Contras yields, ceteris paribus, an average of 8.98 percentage points in increased support, while use of a balanced format creates a slightly larger effect in the opposite direction. The pollsters' use of ideological labels increases support by an estimated 4.38 percentage points. Conversely, mentioning the amount of money to be spent on the policy tends to cause support for the president's policy to drop by an estimated 3.94 percentage points.[8]

In conclusion, the results of this analysis confirm the traditional wisdom but also offer reason for optimism. The way in which a question is posed can alter the results of a poll considerably.[9] Fortunately, however, question wording has potentially distinguishable effects. Rather than throwing up our hands in despair when presented with seemingly noncomparable poll results, we should endeavor to determine systematically what the actual effects of these differing methodologies are, for then it is possible to make more valid comparisons across polls.

---

[6] The critical value for significance at the .05 level for 21 degrees of freedom is 1.721; the t ratio for the "dollar amount" variable is 1.69.

[7] Because the number of data points is not large and the number of different combinations of question wording cues is limited, there is justifiable concern about multicollinearity among the independent variables in our model, which might prevent us from making accurate assessments of their impact on Contra support. The equation in Table 1 shows no obvious signs of multicollinearity; for example, there is no discrepancy between the overall fit of the model and the significance of the individual variables, and there are no coefficients with perverse signs or unexpectedly small magnitudes (see Agresti and Agresti, 1979:390–393). We decided nevertheless to examine the bivariate Pearson's correlations among the independent variables, and these are presented in Appendix 2. The correlation matrix shows one bivariate relationship large enough to present a possible problem: the correlation of .83 between "ideological label" and "balanced format." This relationship results from the fact that the balanced questions in our sample *always* used an ideological label, although a labeling term appeared in two nonbalanced questions as well (see Appendix 1). Appendix 3 shows the effects of eliminating from our multivariate model each one of these two variables—"balanced format" and "ideological label"—in turn. Just as one would expect, because "balanced format" and "ideological label" push the dependent variable in opposite directions, removing the controlling effects of the excluded variable weakens the estimated effect of the remaining variable; "balanced format" holds up better, but we suspected from Table 1 that its impact is stronger anyway. Until more data are assembled that can differentiate these influences more cleanly, the reader can choose which is the greater evil: underspecification or multicollinearity.

[8] The Durbin-Watson test statistic for this model is in the "region of indeterminacy." There is no evidence, therefore, that statistically significant autocorrelation is affecting our estimates.

---

[9] It may be asked whether the poll results vary because of "house effects" above and beyond the different question wordings utilized by the different polling organizations. A test of this hypothesis was attempted by adding dummy variables representing the different polling organizations to the model in Table 1. The model could not withstand this many variables, so the house effects were then assessed one at a time; for example, a dummy variable for CBS/*New York Times* was included in the Table 1 model, then a dummy for ABC/*Washington Post* was substituted, etc. None of these dummy variables proved statistically significant when added to the Table 1 model, suggesting that whatever "house effects" might be present are due primarily to question wording. The dummy variable for "Gallup" (coded 1 for polls 7 and 17 in Appendix 1) was a perfect linear function of the other predictors in Table 1, so its impact could not be measured.

Appendix 1. Texts and Codings of Questions on Contra Aid, 1983–1986

Poll	Text and Coding	For	Against	DK/NS
11–15 Jun 83 ABC/*WP*	Do you think the United States should secretly try to overthrow the government of Nicaragua, or not? [Reagan mention = 0, Contra mention = 0, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	13%	78	10
20–26 Jun 83 CBS/*NYT*	The Reagan administration says the U.S. should help people in Nicaragua who are trying to overthrow the pro-Soviet government there. Other people say that even if the country doesn't like the government in Nicaragua, we should not help overthrow it. Do you think we should help the people trying to overthrow the government of Nicaragua, or not? [Reagan mention = 1, Contra mention = 1, Ideological label = 1, Dollar amount = 0, Balanced format = 1]	23%	53	24
28 Jul–1 Aug 83 ABC/*WP*	As you may know, the United States, through the CIA, is supporting the rebels. Would you say you approve of the United States being involved in trying to overthrow the government of Nicaragua? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	20%	62	17
26 Oct 83 CBS/*NYT*	Same wording as 20–26 Jun 83	23%	58	19
27 Oct 83 CBS/*NYT*	Same wording as 20–26 Jun 83	21%	60	19
3–7 Nov 83 ABC/*WP*	Same wording as 28 Jul–1 Aug 83	30%	48	22

18–21 Nov 83 Gallup	The U.S. is providing military assistance to the forces trying to overthrow the leftist government in Nicaragua. Do you approve or disapprove of this U.S. policy? [Reagan mention = 0, Contra mention = 1, Ideological label = 1, Dollar amount = 0, Balanced format = 0]	33%	43	24
12–17 Jan 84 ABC/*WP*	Same wording as 28 Jul–1 Aug 83	23%	55	22
23–26 Apr 84 CBS/*NYT*	Same wording as 20–26 Jun 83	27%	55	18
14–17 Oct 84 CBS/*NYT*	Same wording as 20–26 Jun 83	29%	50	21
23–25 Oct 84 CBS/*NYT*	Do you think the United States government should provide military assistance to the people trying to overthrow the government of Nicaragua, or not? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	30%	44	27
22–26 Feb 85 ABC/*WP*	Should the U.S. be involved in trying to overthrow the government in Nicaragua, or not? [Reagan mention = 0, Contra mention = 0, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	18%	70	12
21–25 Mar 85 ABC/*WP*	Same wording as 22–26 Feb 85	16%	72	13
29 May–2 Jun 85 CBS/*NYT*	Same wording as 20–26 June 83	32%	53	15
29 May–2 Jun 85 CBS/*NYT*	Should we send military supplies and weapons to the people trying to overthrow the government of Nicaragua? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	24%	66	11

Appendix 1. (*Continued*)

Poll	Text and Coding	For	Against	DK/NS
13–22 Jun 85 ABC/*WP*	Same wording as 22–26 Feb 85 —	18%	71	11
27–28 Aug 85 Gallup/*Newsweek*	Should the U.S. be giving assistance to the guerrilla forces now opposing the Marxist government in Nicaragua? [Reagan mention = 0, Contra mention = 1, Ideological label = 1, Dollar amount = 0, Balanced format = 0]	29%	58	13
6 Mar 86 ABC	President Reagan is asking Congress for new military aid for the Nicaraguan rebels known as the "Contras." Do you agree or disagree with Reagan that Congress should approve the money? [Reagan mention = 1, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	34%	59	8
6 Mar 86 ABC	As you may know the United States is supporting the rebels. Would you say you approve or disapprove of the United States being involved in trying to overthrow the government in Nicaragua? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	32%	52	15
16 Mar 86 ABC	Same wording as second question for 6 Mar 86	30%	54	16
20–24 Mar 86 ABC/*WP*	Same wording as 22–26 Feb 85	28%	62	10
25 Mar 86 ABC	Same wording as second question for 6 Mar 86	42%	53	5

6–10 Apr 86 CBS/*NYT*	Do you think the U.S. should give $100 million in military and other aid to the Nicaraguan rebels known as the *contras*? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 1, Balanced format = 0]	25%	62	13
13–15 Apr 86 NBC/*WSJ*	The Reagan administration has proposed giving $100 million in military, medical, and economic aid to the rebels fighting the Sandinista government in Nicaragua. Do you favor or oppose this proposal? [Reagan mention = 1, Contra mention = 1, Ideological label = 0, Dollar amount = 1, Balanced format = 0]	33%	55	12
28–29 Apr 86 NBC/*WSJ*	Same wording as 6–10 Apr 86	33%	54	13
15–17 Apr 86 YCS	Do you favor or oppose U.S. efforts to help the rebels in Nicaragua fight against the government troops in that country? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	34%	41	26
24–28 Apr 86 ABC/*WP*	Do you generally favor or oppose the U.S. granting $100 million in military and other aid to the Nicaraguan rebels known as the *contras*? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 1, Balanced format = 0]	28%	65	7
19–24 Jun 86 ABC/*WP*	Do you generally favor or oppose the U.S. granting military and other aid to the Nicaraguan rebels known as the *contras*? [Reagan mention = 0, Contra mention = 1, Ideological label = 0, Dollar amount = 0, Balanced format = 0]	29%	62	8

Notes: ABC = ABC News; ABC/*WP* = ABC News/*Washington Post;* CBS/*NYT* = CBS News/*New York Times;* NBC/*WSJ* = NBC News/*Wall Street Journal;* YCS = Yankelovich Clancy Shulman.

Appendix 2. Bivariate Pearson's Correlation Matrix

	Percent Support	Time Trend	Contra Mention	Ideological Label	Dollar Amount	Balanced Format	Reagan Mention
Percent support	1.0	.53	.62	.01	.17	−.10	.34
Time trend		1.0	.04	−.47	.43	−.45	−.01
Contra mention			1.0	.30	.19	.24	.38
Ideological label				1.0	−.26	.83	.46
Dollar amount					1.0	−.21	.09
Balanced format						1.0	.65
Reagan mention							1.0

Appendix 3. Alternative Models

Variable	Coefficient	Std. Error	*T* Ratio	Beta
Without "balanced format"				
Contra mention	9.62	2.26	4.25***	.58
Time trend	.01	.002	3.97***	.59
Ideological label	−.12	2.34	−.05	−.01
Dollar amount	−3.75	2.53	−1.49*	−.20
Reagan mention	1.97	1.91	1.03	.15
Constant	12.47	2.39	5.23***	—
	R^2 = .68	Adjusted R^2 = .61	N = 28	
Without "ideological label"				
Contra mention	9.80	2.08	4.70***	.59
Time trend	.008	.002	3.28***	.48
Dollar amount	−4.33	2.38	−1.82**	−.24
Reagan mention	4.43	2.24	1.98**	.34
Balanced format	−4.65	2.89	−1.61*	−.30
Constant	13.65	2.29	5.96***	—
	R^2 = .72	Adjusted R^2 = .65	N = 28	

*$p < .20$, two-tailed. **$p < .10$, two-tailed. ***$p < .01$, two-tailed.

References

Agresti, Alan, and Barbara Finlay Agresti (1979)
Statistical Methods in the Social Sciences. San Francisco: Dellen.

Bishop, George F. (1987)
"Experiments with the middle category in survey research." Public Opinion Quarterly 51:220–232.

Bishop, George F., Alfred Tuchfarber, and Robert W. Oldendick (1978)
"Change in the structure of American political attitudes: The nagging question of question wording." American Journal of Political Science 22:250–269.

Borrelli, Stephen, Brad Lockerbie, and Richard G. Niemi (1987)
"Why the Democratic–Republican partisanship gap varies from poll to poll." Public Opinion Quarterly 51:115–119.

Congressional Quarterly Almanac (1985)
"Congress votes to resume Nicaraguan rebel aid." Washington: Congressional Quarterly Press.

Kenworthy, Eldon (1987)
"Selling the policy." Pp. 159–181 in Thomas Walker (ed.), Reagan vs. the Sandinistas: The Undeclared War on Nicaragua. Boulder, CO: Westview Press.

Mueller, John (1973)
War, Presidents, and Public Opinion. New York: Wiley.

Rosen, Cory M. (1973)
"A test of presidential leadership of public opinion: The split ballot technique." Polity 6:282–290.

Speakes, Larry (1984)
"Statement to the press." Pp. 517–519 in Weekly Compilation of Presidential Documents, 16 April. Washington: U.S. Government Printing Office.
Thomas, Dan, and Lee Sigelman (1985)
"Presidential identification and policy leadership: Experimental evidence on the Reagan case." In George Edwards, Steven Shull, and Norman Thomas (eds.), The Presidency and Public Policy Making. Pittsburgh: University of Pittsburgh Press.
Yeric, Jerry L., and John R. Todd (1983)
Public Opinion: The Visible Politics. Itasca, IL: F. E. Peacock.

RACE DIFFERENCES IN ABORTION ATTITUDES: SOME ADDITIONAL EVIDENCE

CLYDE WILCOX

Abstract Although Combs and Welch reported a trend of decreasing racial differences in abortion attitudes, Hall and Ferree used data from the 1982 General Social Survey to argue that racial differences were not declining. This paper updates this debate through the 1988 General Social Survey and concludes that racial differences have indeed declined over time. Moreover, when new religious items introduced in the 1984 survey are included in the multivariate analysis, blacks are not significantly different from whites in their support of legal abortion. This finding obscures a more interesting pattern, however, of offsetting, statistically significant racial differences among respondents of the same gender—black men are significantly less supportive of abortion than white men, and black women are significantly more supportive than white women.

Social scientists have consistently reported that blacks are less supportive of legal abortion than whites, even after controls for socioeconomic status and other demographic variables, gender-role attitudes, attitudes toward sexual behavior, and religion (Combs and Welch, 1982; Hall and Ferree, 1986; Secret, 1987). These differences are surprising in light of differences in abortion rates: black women are twice as likely to have abortions as whites, a ratio which has held relatively constant for more than a decade (Henshaw and Silverman, 1988).[1]

Combs and Welch (1982) used data from the National Opinion Research Center's General Social Survey (GSS) and found a lessening of

CLYDE WILCOX is an Assistant Professor of Government at Georgetown University. The author thanks Elizabeth Cook and Ted Jelen for helpful comments. The data were made available by the Inter-university Consortium for Political and Social Research. All interpretations are the author's.

[1] The rate for nonwhite women in 1988 was 5.3 per 100; for whites it was 2.3. In 1977 the figures were 5.9 and 2.1. These figures would suggest that the rate of abortion for blacks may have fallen somewhat in the late 1970s and early 1980s.

Public Opinion Quarterly Volume 54:248–255 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-02/$2.50

racial differences between 1972 and 1980. They reported that controls for demographic variables and religion substantially reduced racial differences, indicating they were largely due to differences in socioeconomic status and religious attachment. Using data from the 1982 GSS, which oversampled blacks, Hall and Ferree (1986), however, found no trend toward declining racial differences. The authors argued that in addition to demographic and religious variables, three sets of attitudes are important in understanding racial differences: attitudes toward gender roles, attitudes toward ideal family size, and attitudes toward premarital sex. The addition of these attitudinal measures to the demographic and religious variables did not reduce racial differences in abortion support, but it did result in a substantial increase in the proportion of variance explained.

Since the Hall and Ferree analysis, the GSS has been conducted several times, including another oversample of blacks in 1987. New religious variables have been added to the GSS, including a measure of personal religiosity (frequency of prayer), a measure of doctrine (interpretation of the Bible), and a measure of denominational fundamentalism. This note updates the Combs and Welch and Hall and Ferree analyses through 1988 to determine trends in racial differences, assess the impact of the new religious items in explaining racial differences, and explore racial differences among respondents of the same gender.

Method

The dependent variable is the same as that used by Combs and Welch and by Hall and Ferree—an additive index of six abortion items, with values ranging from 0, when abortion was opposed in every case, to 6, where approval was given for abortion in every case. The same demographic variables are also employed: income, education, gender, age, Southern residence presently and at age 16, urban residence currently and at age 16. Hall and Ferree's attitudinal measures, which were not included in the 1987 survey, are used for one portion of the analysis.

Combs and Welch reported that religious variables explained a good deal of the observed racial differences, and it is possible that better measures would further reduce this gap. The 1984 survey introduced two new religious variables: (1) frequency of prayer, which I have combined with frequency of attendance and strength of attachment to denomination (RELITEN) for a measure of religiosity, and (2) a measure tapping interpretation of the Bible. In 1987 NORC introduced a measure of denominational fundamentalism, FUND, which represents a coding scheme for denomination and is available for all earlier surveys.[2] It has been used in place of the dummy variable in the Combs and Welch and Hall and Ferree studies. A dummy variable was created to identify Catholics.

Trends in Abortion Attitudes

Hall and Ferree argued that Combs and Welch had overestimated the decline in racial differences in abortion by their choice in combining surveys: the 1978 survey contained the lowest differences observed to date, and when combined with the 1980 survey it served to lower the mean. With a few more years of data, it seems that the 1982 survey was more atypical—the 1982 racial gap was the highest for the 1974–88 period. Table 1 presents the mean differences for each year between 1972 and 1988. Note that the 1982 survey stands out against a pattern of generally decreasing racial differences. The decline is not sharp or steady, but the last five surveys provided four of the five smallest aggregate differences in racial attitudes, and the first two surveys provided the largest aggregate differences. The decrease in racial differences is due primarily to declining support among whites for legal abortion; black support remains relatively constant throughout the period.

Factor analysis suggests that abortion attitudes can be understood as composed of two related attitudes: attitudes toward abortion under traumatic circumstances, such as rape, health problems, or fetal deformity, and attitudes toward elective abortion when the mother is poor, single, or wants no more children (see Jelen, 1984, for a similar result). Additive scales were constructed counting the number of prochoice responses to these two sets of items. Table 1 presents the results of this analysis over time. Note that although there is at most a small trend for differences on elective abortion to decline, there is a clear pattern of decreasing racial differences on attitudes toward traumatic abortion. Black attitudes have remained fairly constant toward traumatic abortion over time; white support declined after 1982.

To test the statistical significance of the interactions between year and race, regression equations were estimated with year, race, and an

---

2. Jelen (1989) has demonstrated that the two variants of the Bible item, BIBLE and BIBLEY, are interchangeable for most purposes. This measure may be thought of as a doctrinal measure of fundamentalism, while FUND is a denominational measure. FUND is a three-point scale, with conservative denominations coded at 3, moderate denominations at 2, and liberal denominations at 1. Although there are a few problems with the FUND coding scheme, I have elected to use this item rather than creating a dummy variable because it is easily available to future researchers. For both blacks and whites, the Bible question is correlated with FUND, but the correlation is higher for whites (.34) than for blacks (.14).

Table 1. Abortion Attitudes and Race, 1972–87

	Blacks	Whites	Difference
Abortion attitudes			
1972	2.89	3.76	.89
1973	3.24	4.13	.89
1974	3.43	4.13	.70
1975	3.43	4.01	.58
1976	3.23	4.06	.83
1977	3.33	4.08	.75
1978	3.27	3.81	.54
1980	3.38	3.99	.61
1982	3.25	4.06	.81
1983	3.22	3.67	.45
1984	3.13	3.84	.71
1985	3.34	3.69	.35
1987	3.24	3.79	.55
1988	3.37	3.65	.28
Traumatic abortion			
1972	1.88	2.45	.57
1973	2.17	2.61	.44
1974	2.31	2.62	.31
1975	2.18	2.56	.38
1976	2.13	2.57	.44
1977	2.11	2.60	.49
1978	2.18	2.55	.37
1980	2.18	2.54	.36
1982	2.19	2.60	.41
1983	2.21	2.47	.26
1984	2.13	2.49	.36
1985	2.25	2.45	.20
1987	2.22	2.46	.24
1988	2.24	2.44	.20
Elective abortion			
1972	1.02	1.31	.29
1973	1.06	1.52	.46
1974	1.12	1.51	.39
1975	1.24	1.45	.21
1976	1.10	1.48	.38
1977	1.22	1.48	.26
1978	1.10	1.26	.16
1980	1.20	1.45	.25
1982	1.06	1.46	.40
1983	1.02	1.20	.18
1984	.98	1.35	.37
1985	1.08	1.24	.16
1987	1.01	1.33	.32
1988	1.13	1.21	.08

NOTES: Entries are mean number of responses supportive of legal abortion. Respondents missing four or more items have been excluded from the analysis. There was no GSS in 1979 or 1981, and the abortion items were not asked in the 1986 survey.

interaction term as independent variables. The interaction term was significant for the traumatic abortion ($T = 5.0$) and for the inclusive abortion scale ($T = 3.5$), but not for elective abortion ($T = 1.65$, $p \leq .10$).

Multivariate Analysis

Both Combs and Welch and Hall and Ferree found that controls for demographic and religious variables substantially reduced racial differences in abortion attitudes. The data in Table 2 also support this conclusion. For each period, controls for demographic and religious variables substantially reduce the regression coefficient for race—in 1987–88 these controls reduce the coefficient to nonsignificance.

For 1984–85 and 1987–88, I have estimated a separate equation which includes the new religious variables—the expanded religiosity scale and interpretation of the Bible. Inclusion of the new religious items resulted in an increase in the proportion of variance explained. Race was not a significant predictor of abortion attitudes in either of the equations that included the new religious variables.

Unfortunately, the 1987 survey did not include the attitudinal items used by Hall and Ferree. This raises the possibility that the disappearance of racial differences in 1987 is due to model misspecification—inclusion of attitudinal variables might restore race to its predictive significance. Additional analysis using the 1988 data suggested that this was not the case. Details of this analysis may be obtained from the author.

Race and Gender

There is also reason to suspect that there is an interaction between race and gender. Ransford and Miller (1983) reported that black men were more traditional in gender-role attitudes than white men, but found no significant racial differences among women. Petchesky's (1984) analysis also suggests the possibility of gender–race interaction on abortion attitudes. She notes that "abortion has always been for black women an important 'fact of life,' a well known option practiced by them despite its illegality, despite moral and religious opposition to it in the black community, and despite its association with genocide by black male political figures" (p. 154). This argument suggests that black men may be less supportive of legal abortion than white men for political reasons, while black women may be more supportive than white women because of the greater frequency of abortion in the black com-

Table 2. Race, Abortion Attitudes, and Multivariate Controls

	Race	Race Demog	Race Demog Relig	Race Demog Relig New Relig
Race				
1972–74				
B	–.90*	–.58*	–.18*	
Beta	–.15	–.10	–.03	
R^2	.02	.09	.14	
1975–77				
B	–.70*	–.40*	–.30*	
Beta	–.11	–.06	–.05	
R^2	.01	.09	.16	
1978–80				
B	–.57*	–.43*	–.29*	
Beta	–.09	–.06	–.04	
R^2	.01	.09	.16	
1982				
B	–.83*	–.67*	–.45*	
Beta	–.13	–.10	–.07	
R^2	.02	.13	.18	
1983–85				
B	–.56*	–.44*	–.18*	–.16[a]
Beta	–.08	–.07	–.03	–.02
R^2	.01	.08	.16	.20
1987–88				
B	–.60*	–.40*	–.12	–.08
Beta	–.10	–.07	–.02	–.01
R^2	.00	.10	.16	.20

NOTES: Unstandardized and standardized regression coefficients for race with and without controls, and R^2 for each equation for each year. Demographic controls included income, education, age, gender, region, region at age 16, size of community, and size of community at age 16. Religious variables included Catholicism, fundamentalism, and religiosity (frequency of attendance and commitment to denomination). New religious variables added frequency of prayer to the religiosity scale, and interpretation of Bible as separate item. Equations for 1982 and 1987 were weighted to reflect population distribution on race.

* Statistically significant.

[a] Equation included data from 1984–85 only.

munity. Hall and Ferree reported a gender gap among blacks in 1982, with men significantly less supportive of legal abortion.

Separate equations predicting support for elective and traumatic abortion among men and women were estimated to examine the gender–race interaction (not shown). For elective abortion, offsetting racial gaps emerged in the 1987–88 data, with black women significantly more supportive than white women and black men significantly less supportive than white men. The gender–race interaction was statistically significant.[3] Black men were less supportive than whites of traumatic abortion, but race was not a significant predictor of women's attitudes.

These data offer no clear explanations of these differing racial gender gaps. The race–gender interaction is first significant in the 1987 data, and increases in 1988, but the black gender gap existed in earlier surveys (Hall and Ferree, 1986). The GSS does not contain items to directly test whether racial differences among women may be due in part to higher rates of performed abortions among black women, nor can I test Petchesky's theory of black gender differences, since no items on perceived genocide appear in the survey. Separate analysis (not shown) suggests that neither perceptions of racial discrimination nor black consciousness is a significant predictor of abortion attitudes among black men.

Conclusions

This study confirms Combs and Welch's finding that racial differences on abortion attitudes are declining. The 1982 survey is an exception to the general trend, but the racial gap has gradually eroded over the course of the 17 years covered by the General Social Surveys. The introduction of new religious variables in the later surveys help explain away racial differences, suggesting that they are in part a function of the greater religiosity and doctrinal orthodoxy of the black population. Although these results support Combs and Welch's conclusion of a narrowing racial gap, there is less support for their explanation for this narrowing. Blacks are not becoming more supportive over time as their socioeconomic position improves—whites are becoming less supportive as the political climate changes.

In 1987–88, black women were more supportive of elective abortion than their white counterparts after multivariate controls, while this pattern was reversed for men. The data permit no direct test of Petchesky's suggestion that abortion is associated with genocide for black males, but her argument that abortion is perceived as necessary by black women seems a reasonable explanation for racial differences among women.

---

[3] *Z* scores were computed by a formula where the numerator is the difference of unstandardized regression coefficients and the denominator is the square root of the sum of the squared standard errors of those coefficients.

More research is needed, using better measures of black family structure, culture, and perceptions of the politics of abortion. Luker (1984) provided rich insight into the worldviews of abortion activists through in-depth interviews. A similar technique might be useful to understand racial differences in abortion support.

References

Combs, M., and S. Welch (1982)
"Blacks, whites, and attitudes toward abortion." Public Opinion Quarterly 46:510–520.
Hall, E., and M. Ferree (1986)
"Race differences in abortion attitudes." Public Opinion Quarterly 50:193–207.
Henshaw, S., and J. Silverman (1988)
"The characteristics and prior contraceptive use of U.S. abortion patients." Family Planning Perspectives 20(4):158–168.
Jelen, T. (1984)
"Respect for life, sexual morality, and opposition to abortion." Review of Religious Research 25:220–231.
——— (1989)
"Biblical literalism and inerrancy: Does the difference make a difference?" Sociological Analysis 49:421–429.
Luker, K. (1984)
Abortion and the Politics of Motherhood. Berkeley: University of California Press.
Petchesky, R. (1984)
Abortion and Woman's Choice. New York: Longman.
Ransford, E., and J. Miller (1983)
"Race, sex, and feminist outlooks." American Sociological Review 48:46–59.
Secret, P. (1987)
"The impact of region on racial differences in attitudes toward legal abortion." Journal of Black Studies 17:347–369.

THE POLLS—A REPORT

NATIONAL SERVICE

TOM W. SMITH

> Ask not what your country can do for you—ask what you can do for your country.
>
> John F. Kennedy, January 1961

Recently there has been a flurry of interest in the idea of national service (Kuntz, 1989; Shapiro, 1989; Sheler and Whitman, 1989; Moskos, 1988). Eight different legislative proposals were introduced in 1989 calling for some form of national service. These, as well as other national service programs, are quite varied in their goals and forms.

Proponents of national service for youth find many merits in the program. Some see national service as engendering a sense of community among self-centered and atomistic youth. As Moskos (1988) noted "shared civic duties become the social glue that holds society together," while Janowitz (1983) goes further to see it as promoting "the reconstruction of patriotism." Others see national service as principally a program to help youths, partly an economic program to reduce youth unemployment and teach job skills, and partly a moral program to teach self-discipline and responsibility (Moore, 1989). Some others emphasize the public services and alleviation of social problems, how others such as the elderly or the environment will benefit from the exertions of the young. Still others believe that the invigoration of volunteerism is as much a plus as the good works themselves. Of course many mention some, or even all, of the above claimed advantages (James, 1910; Moskos, 1988; Boorstin, 1989; Smith, 1989).

Depending on what goals are emphasized, the form of the proposed national service programs varies greatly. Some involve voluntary programs, and these sometimes focus on assistance to disadvantaged youths, while others stress service by advantaged youths. For the former the model is Roosevelt's Civilian Conservation Corps (CCC), while for the latter Kennedy's Peace Corps and Johnson's VISTA

TOM W. SMITH is Director of the General Social Survey at the National Opinion Research Center, University of Chicago.

Public Opinion Quarterly Volume 54:273–285 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-02/$2.50

(Volunteers in Service to America) are prototypes. Others call for mandatory involvement in either military or civilian national service.

Public response to the idea of national service varies according to the form and goal of the program. Youth corps programs to help the disadvantaged have been popular since Roosevelt's CCC. Large majorities backed the CCC in the 1930s, and in the 1960s and 1970s generally 80% of the public supported the return of the CCC (see "FDR's CCC" and "Youth Corps: Voluntary"). Mandatory youth corps plans have received less support than the voluntary CCC plans, but in the 1960s and early 1970s majorities consistently approved of such proposals (see "Youth Corps: Mandatory").

In the 1970s attention mainly shifted from youth corps plans for the economically disadvantaged to general national service programs that offered a choice of military or civilian participation. As in the case of the youth corps plans, voluntary national service proposals have been quite popular. Support among the young (18–24) rose from 77% in 1979 to 87% in 1987,[1] while like numbers of all ages supported voluntary national service (see "National Service: Voluntary").

Mandatory national service plans for young men have generally gathered somewhat less, but still majority, support (see "National Service for Men: Mandatory"). Support has varied over time however, falling from 1969 to 1979, rising in the early 1980s, and apparently dropping again in the late 1980s. The rise in the early 1980s and subsequent decline also appears among teenagers.

Mandatory service for young women has received less support than for men (see "National Service for Women: Mandatory"). Usually small majorities have opposed such plans, and support has not varied much over time.

However, there has been a decline over time in the greater support for national service for men than for women (see "Support for Mandatory National Service for Men and Women"). While approval of mandatory national service for men exceeded approval of national service for women by 34 percentage points in 1969, this edge steadily decreased to 9–11 percentage points in 1982–87. Similarly, among teenagers the difference in approval for men vs. women serving declined between 1977 and 1988.

Differences in support for mandatory national service for men by age groups have widened considerably over time (see "Support for National Service by Age"). In the 1960s there were little consistent age differences. Since then those over 50 have been more supportive of national service than those under 30; the difference grew to 27 percentage points in 1987. This change seems to reflect changes in the perceived self-interest of the young. National service, of course, is an obligation borne by the young, and self-interest theory suggests one would find less support among those at risk than those beyond the targeted ages. This is the age pattern that prevails since 1973 for national service for men and since 1969 for national service for women. The expected pattern does not appear for men in 1966 and 1969, however, almost certainly because the military draft for men was in force until early 1973. In those years, the obligation of national service with the possibility of nonmilitary service actually offered greater choice for young men than the system in place. The role of self-interest is further shown by the fact that there is no age difference in support for voluntary national service for men and women. In brief, due to self-interest, the young have opposed mandatory service more than the old, except in the 1960s when national service offered a possible alternative to the military draft for young men. Voluntary service, which offers an opportunity rather than obligation, is highly supported by all age groups.

---

[1] Because of the small subgroup size, this difference is not statistically significant, however.

Mandatory national service proposals offer a choice between military and nonmilitary service. The general public favored military service over nonmilitary service for men in 1966, and split evenly in 1979. In 1979 the general public favored civilian service over military service for women. Similarly, among teenagers in 1986 and 1988 more personally favored nonmilitary service than military service. In addition, the young have not favored military service as much as all adults have (see "Military vs. Nonmilitary Service").

Both voluntary and mandatory national service programs would depend on the cooperation of the young for success. Substantial minorities of young people have expressed an interest in performing national service. In 1979 15% of those 18–24 reported that they were definitely interested in volunteering for national service, and another 20% expressed some interest. In 1982 among those 18–29 43% had at least some interest. By other indicators, however, youth interest appears to have been slipping. Among first year college students, the portion thinking it was essential or very important to "participate in an organization like the Peace Corps or VISTA" dropped from 21% in 1966 to 7.2% in 1988. While part of this decline comes from the ending of the military draft and probably a decline in the visibility of these programs (AHF Market Research, 1983), it was similar to a decline from 29% in 1970 to 18.5% in 1985 of those saying it was essential or very important to "participate in a community action program."[2]

[2] Whether this indicates a societal decline in either altruism or volunteerism is doubtful. On one hand, participation in charitable activities rose from 1977 to 1987. On the other hand, membership in voluntary associations has been steady from 1973 to 1989 (Smith, forthcoming).

One apparent way to kindle enthusiasm in national service is to appeal to self-interest. When educational scholarships are given for national service, as in the Mikulski and Nunn/McCurdy plans, over 80% of teenagers in 1989 were very or somewhat interested, and their interest rose notably as the amount of credits grew (Bezilla, 1989).

Large majorities have consistently supported voluntary national service in general and youth corps plans for the disadvantaged. Smaller majorities have backed mandatory national service for men and obligatory participation in a youth corps for the disadvantaged. Mandatory national service for women has usually been opposed by a majority, but the gap between public support for service by women vs. men has declined over time.

While Kennedy's altruistic appeal to serve one's country undoubtedly is a source of support for national service among the public at large and a motivation for personal involvement among the young, self-interest is clearly a major factor in shaping the attitudes of the young to national service. Either the stick of the military draft or the carrot of educational scholarships make national service more appealing to the young.

CIRP—Cooperative Institutional Research Program, Higher Education Research Institute, University of California at Los Angeles
GALLUP—Gallup Organization
GALLUP/YS—Gallup Youth Surveys, George H. Gallup International Institute
NORC/GSS—National Opinion Research Center, General Social Survey
ROPER—Roper Organization

References

AHF Marketing Research (1983)
"A study of public attitudes toward and knowledge about the Peace Corps." Unpublished report.
Bezilla, Robert (1989)
"Surveys of the attitudes of teen-agers towards national service proposals." Unpublished report, George H. Gallup International Institute.
Boorstin, Daniel J. (1989)
"Can patriotism be legislated?" U.S. News and World Report, 13 February, 26.
James, William (1910) [1971]
"The moral equivalent of war." In John K. Roth (ed.), The Moral Equivalent of War and Other Essays. New York: Harper.
Janowitz, Morris (1983)
The Reconstruction of Patriotism. Chicago: University of Chicago Press.
Kuntz, Phil (1989)
"Nunn-McCurdy plan ignites national service debate." Congressional Quarterly Weekly Report 47 (March 25):645–648.
Moore, Thomas (1989)
"For whom the corps toils." U.S. News and World Report, 13 February, 24–25.

Moskos, Charles C. (1988)
A Call to Civic Service: National Service for Country and Community. New York: The Free Press.
Shapiro, W. (1989)
"The gap between will and wallet." Time, 6 February, 32.
Sheler, Jeffery L., and David Whitman (1989)
"The push for national service." U.S. News and World Report, 13 February, 21–23.
Smith, Tom W. (1989)
"Random probes of GSS questions." International Journal of Public Opinion Research 1:305–325.
——— (forthcoming)
"Trends in voluntary group membership: Comments on Baumgartner and Walker." American Journal of Political Science.

FDR's CCC

GALLUP: Are you in favor of the CCC camps?

	1936
Yes	82%

GALLUP: Do you think CCC camps should be made permanent?

	1938
Yes	78%

GALLUP: Do you think spending should be reduced on farm benefits; on public works not connected with defense; on WPA; on the CCC; on the National Youth Administration (NYA)?

		1941			
	Farm Benefits	*Public Works*	*WPA*	*CCC*	*NYA*
Yes	36%	66%	62%	40%	37%
No	50	27	32	51	47
No opinion	14	7	6	9	16
	100	100	100	100	100

Youth Corps

VOLUNTARY

GALLUP: It is proposed that the federal government set up youth camps—such as the CCC (Civilian Conservation Corps) camps of the 1930s—for young men 16 to 22 years old who want to learn a trade and earn a little money by outdoor work. Do you think this is a good idea or a poor idea?

	1961	*1962*	*1963*[a]	*1976*
Good idea	80%	79%	89%	85%
Poor idea	13	16	6	10
No opinion	7	5	5	6
	100	100	100	101
N	1,647	1,599	1,588	1,527

[a] Congress is now considering a plan to establish a youth conservation corps for young men who want to learn a trade and earn a little money by outdoor work. Do you think this is a good or a poor idea?

MANDATORY

GALLUP: Some people say that all young men between the ages of 16 and 22 who are out of school and out of work should be required to join a Youth Conservation Corps to carry on their education, learn a trade, and earn a little money. Do you approve of this plan?

	1961[a]	*1962*[b]	*1964*	*1968*[c]	*1971*
Approve	59%	51%	73%	73%	66%
Disapprove	34	43	21	23	27
No opinion	7	6	5	4	7
	100	100	99	100	100
N	1,647	1,599	1,634	1,507	1,639

[a] Do you think young men who are not in school, do not have a job, and are not learning a trade should be required to go to one of these (youth) camps (such as the Civilian Conservation Corps camps of the 1930s), or not?

[b] Do you think young men who are not in school, do not have jobs, and are not learning a trade should be required to go to one of these camps (as the CCC [Civilian Conservation Corps] camps of the 1930s), or not?

[c] Some people say that all young men between the ages of 16 and 22, who are out of school and out of work and not in the army, should be required to join a youth conservation corps to carry on their education, learn a trade, and earn a little money. Do you approve or disapprove?

National Service

VOLUNTARY

GALLUP: Would you favor or oppose a system of voluntary national service in which young people (both men and women), after high school or college, would be given opportunities to serve one year, either in the military forces or in non-military work here or abroad, such as VISTA or the Peace Corp?

		18–24		*All Ages*	
	1979	*1982*	*1987*[a]	*1982*	*1987*
Favor	77%	80%	87%	77%	83%
Oppose	14	15	6	15	11
No opinion	9	5	7	8	6
	100	100	100	100	100
N	197		132	1,514	1,549

[a] Omits "such as VISTA or the Peace Corps."

MANDATORY FOR UNEMPLOYED

GALLUP: As you may know, the United States has a youth unemployment problem. It has been suggested that we develop a national youth service which would require every young man under the age of 20 who is unemployed, not attending school or college, to take vocational or on-the-job training, or to perform public or military service until he reaches the age of 20. Would you approve or disapprove of such a national youth plan for young men?
Would you approve or disapprove of such a national youth service plan for young women?

	1979	
	Men	*Women*
Approve	67%	62%
Disapprove	27	31
Don't know	7	7
	101	100
N	1,511	

MANDATORY, MEN AND WOMEN

ROPER: One proposal before Congress would require all young Americans—both men and women—to serve for one year in a national service program. They could serve in the armed forces, or do some other form of service such as work in a hospital, or provide help for the elderly, or work in day care centers for children, etc. There would be special incentives for those choosing military service. Would you be in favor of or opposed to one year of compulsory national service for all young people at any time of their choice between age 17 and 25?

	1979	*1979*	*1980*	*1981*	*1986*
Approve	54%	57%	62%	62%	50%
Disapprove	38	36	29	33	42
Don't know	8	7	9	5	8
	100	100	100	100	100
N	2,004	3,944	2,001	2,000	1,993

GALLUP: It has been proposed that all young men and women be required to give one year of national service, either in the armed forces or in civilian social work. In return, they would receive government benefits to help cover the cost of their college education. Does this proposal sound like a good idea or a poor idea to you?

	1983
Good idea	64%
Poor idea	30
Don't know	5
	99
N	1,510

National Service for Men: Mandatory

GALLUP: Would you favor or oppose requiring all young men to give one year of service to the nation—either in the military forces, or in non-military work here or abroad, such as VISTA or the Peace Corps?

	Adults							
	1969	*1971*	*1973*	*1976*	*1979*	*1981*[a]	*1982*[a]	*1987*[b]
Favor	78%	68%	64%	62%	60%	71%	71%	55%
Oppose	16	25	29	33	33	24	24	39
Don't know	5	7	7	5	7	5	5	6
	99	100	100	100	100	100	100	100
N	1,461	1,504	1,576	1,559	1,541	1,515	1,511	1,549

	Teenagers			
	1977	*1984*	*1986*	*1988*
Favor	47%	62%	62%	55%
Oppose	46	34	35	43
Don't know	7	4	3	2
	100	100	100	100
N	1,035	504	528	1,003

[a] Would you favor or oppose requiring all young men to give one year of service to the nation—either in the military forces or in non-military work here or abroad, such as work in hospitals or with elderly people?

[b] Would you favor or oppose requiring all young men to give one year of service to the nation, either in military forces or in non-military work, here or abroad, such as VISTA, the Peace Corps, or in a local community or city service program?

GALLUP: Would you favor or oppose requiring all young men to give two years of service—either in the military forces, or in non-military work here or abroad, such as the Peace Corps?

	1966
Favor	72%
Oppose	21
No opinion	7
	100
N	1,562

GALLUP: Congress is now considering a proposal to replace the selective service system—that is, the draft—with a national service system. Under this system, a young man of 18 could choose to do any one of these three things:

A. He could volunteer for military service, or
B. He could volunteer for civilian service—e.g., helping in hospitals, teaching school, working in programs like VISTA, or
C. He could take his chances on being drafted?

How does this plan sound to you? Would you like to have your Congressman vote for or against this proposal?

	1970
Vote for	71%
Vote against	20
No opinion	9
	100
N	1,513

NORC/GSS: How would you feel about a program that required all young men to give one year of service to the nation—either in the military forces or in non-military work such as hospitals or with elderly people—would you strongly favor it, probably favor it, probably oppose it, or strongly oppose it?

	1982	*1984*
Strongly favor	37%	33%
Probably favor	35	37
Probably oppose	15	16
Strongly oppose	9	9
Don't know	2	5
	98	100
N	1,506	1,473

National Service for Women: Mandatory

GALLUP: Would you favor or oppose requiring all young women to give one year of national service to the nation—either in the military forces, or in non-military work here or abroad, such as VISTA or the Peace Corps?

			Adults		
	1969[a]	*1976*	*1979*	*1981*[b]	*1987*[c]
Favor	44%	41%	41%	54%	44%
Oppose	49	51	53	40	50
No opinion	7	9	7	6	6
	100	101	101	100	100
N	1,461	1,559	1,541	1,515	1,547

		Teenagers		
	1977	*1984*	*1986*	*1988*
Favor	38%	51%	56%	50%
Oppose	55	44	42	49
No opinion	7	5	2	1
	100	100	100	100
N	1,035	504	528	1,003

[a] How about young women—would you favor or oppose requiring all young women to give one year of service to the nation—either in nursing or some other form of public help?

[b] Would you favor or oppose requiring all young women to give one year of service to the nation, either in military forces or in non-military work here or abroad, such as VISTA, the Peace Corps, or in a local community or city service program?

[c] Would you favor or oppose requiring all young women to give one year of service to the nation, either in military forces or in non-military work here or aboard, such as VISTA, the Peace Corps, or in a local community or city service program?

NORC/GSS: How would you feel about a program that required all young women to give one year of service to the nation—either in the military forces or in non-military work such as in hospitals or with elderly people—would you strongly favor it, probably favor it, probably oppose it, or strongly oppose it?

	1982	*1984*
Strongly favor	27%	23%
Probably favor	34	38
Probably oppose	22	22
Strongly oppose	15	15
Don't know	3	2
	101	100
N	1,506	1,473

NORC/GSS: Suppose that the costs of such a program (that required young men and/or young women to give a year of service to the nation) made it necessary to increase your taxes by a small amount—for example, 5 percent. Would you strongly favor it, probably favor it, probably oppose it, or strongly oppose it? (Based only those who strongly or probably favored plan for men and/or women)

	1982	*1984*
Strongly favor	21%	19%
Probably favor	40	45
Probably oppose	22	24
Strongly oppose	14	10
Don't know	3	2
	100	100
N	1,114	1,075

Support for Mandatory National Service for Men and Women

	Men	*Women*	*Women − Men*
1969[a] (Gallup)	78%	44%	−34
1976 (Gallup)	62	41	−21
1979 (Gallup)	60	41	−20
1981 (Gallup)	71	54	−17
1982 (NORC/GSS)	71	61	−10
1984 (NORC/GSS)	70	61	− 9
1987 (Gallup)	55	44	−11

[a] In 1969 only the male and female choices are not balanced. Women would be subject only to non-military service.

Support for National Service by Age

	Under 30	*30–49*	*50+*	*50+ − LT30*
A. Mandatory, Men				
1966 (Gallup) (2 yrs.)	68%	74%	72%	+ 4
1969 (Gallup)	80	81	75	− 5
1973 (Gallup)	56	67	68	+12
1979 (Gallup)	47	69	63	+16

1982 (NORC/GSS)	61	77	82	+21
1984 (NORC/GSS)	59	77	81	+22
1987 (Gallup)	39	58	66	+27
B. Mandatory, Women				
1969 (Gallup)	38	40	50	+12
1979 (Gallup)	32	46	42	+10
1982 (NORC/GSS)	53	67	67	+14
1984 (NORC/GSS)	52	66	67	+15
1987 (Gallup)	36	50	43	+ 7
C. Voluntary, Men and Women				
1982	80	80	79	− 1
1987	86	81	83	− 3

Military vs. Non-Military Service

GALLUP: Suppose all young men were required to give two years service, would you rather have a son of yours serve in the armed forces, or in non-military work?

	1966	
	All Ages	*18–29*
Military	47%	44%
Non-military	28	31
Don't know	25	25
	100	100
N	1,562	282

GALLUP: Suppose all young men were required to give one year of service, which would you prefer—military service or non-military service?

	1973	*1979*	*1979*
	18–26	*All*	*18–29*
Military	32%	42%	30%
Non-military	64	42	59
Don't know	4	16	11
	1C	100	100
N		1,541	391

GALLUP: Suppose all young women were required to give one year of service, which would you prefer—military or non-military?

	1979	*1979*
	All	*18–29*
Military	20%	17%
Non-military	65	74
Don't know	15	9
	100	100
N	1,541	391

GALLUP/YS: If all young people were required to this (national service), would you personally prefer military or non-military work?

	Teenagers		
	1984	*1986*	*1988*
Military	36%	39%	40%
Non-military	57	59	59
Don't know	7	2	1
	100	100	100
N	504	528	1,003

Volunteers for National Service

GALLUP: Do you think you, yourself, would be interested in volunteering for such a program (a voluntary national service . . . to serve for one year either in the military forces or in non-military work . . .), or not?

	1979	*1982*
	18–24	*18–29*
Yes, definitely	15%	23%
Yes, might	20	20
No	57	54
No opinion	8	3
	100	100
N	197	

CIRP: Indicate the importance to you personally of each of the following (asked of first year college students):

A. Participate in an organization like the Peace Corps or VISTA

	1966	*1967*	*1968*	*1970*	*1971*	*1972*	*1988*
Essential or very important	21.0%	18.9%	18.4%	19.5%	16.2%	15.8%	7.2%

B. Participate in a community action program

	1970	*1971*	*1972*	*1974*	*1975*	*1976*	*1977*
Essential or very important	29.4%	25.9%	29.1%	27.5%	30.4%	28.8%	29.4%
	1978	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*
Essential or very important	26.7%	26.0%	27.4%	24.1%	22.8%	22.1%	22.2%
	1985	*1986*	*1987*	*1988*			
Essential or very important	22.8%	18.5%	19.8%	21.7%			

NOTE: Based on very large samples; in 1988 *N* = 222,296.

GALLUP: Do you, yourself, happen to be involved in any charity or social service activities, such as helping the poor, the sick, or the elderly?

	4/77	*9/81*	*12/81*	*12/84*	*3/86*	*5/87*	*12/87*
Yes	27%	29%	32%	31%	36%	39%	49%
N	1,540	1,549	1,483	1,505	1,004	503	606

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

ROY PIERCE and PHILIP E. CONVERSE	Attitudinal Sources of Protest Behavior
MARTIN BROWN and ARNOLD POTOSKY	The President's Cancer and the Public's Health
ELIZABETH F. LOFTUS and others	A Tale of Two Questions
JEANNINE M. JAMES and RICHARD BOLSTEIN	Monetary Incentives and Follow-Up Mailings
WILLIAM S. AQUILINO and LEONARD A. LOSCIUTO	Interview Mode Effects in Drug Use Surveys
GABRIEL WEIMANN	The Obsession to Forecast
THE POLLS	Review: The *Times*'s Israeli Poll *by Howard Schuman*
	Report: The Sexual Revolution? *by Tom W. Smith*
COMMENT AND CONTROVERSY	Comments *by Tom W. Smith and Jean-Noël Kapferer*
BOOKS	Reviews and Notes *by Lee Sigelman, Benjamin I. Page, Paul D. Allison, and Norval D. Glenn*

PROCEEDINGS OF THE FORTY-FIFTH ANNUAL CONFERENCE OF THE
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Fall 1990

PUBLIC OPINION QUARTERLY

VOLUME 54 / FALL 1990 / NUMBER 3

Articles

Attitudinal Sources of Protest Behavior in France: Differences Between Before and After Measurement 295
ROY PIERCE and PHILIP E. CONVERSE

The Presidential Effect: The Public Health Response to Media Coverage about Ronald Reagan's Colon Cancer Episode 317
MARTIN BROWN and ARNOLD POTOSKY

A Tale of Two Questions: Benefits of Asking More Than One Question 330
ELIZABETH F. LOFTUS, MARK R. KLINGER, KYLE D. SMITH, and JUDITH FIEDLER

The Effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys 346
JEANNINE M. JAMES and RICHARD BOLSTEIN

Effects of Interview Mode on Self-Reported Drug Use 362
WILLIAM S. AQUILINO and LEONARD A. LO SCIUTO

The Obsession to Forecast: Pre-Election Polls in the Israeli Press 396
GABRIEL WEIMANN

The Polls

Review: The *Times*'s Israeli Poll on the Eve of Shamir's 1989 Visit 409
HOWARD SCHUMAN

Report: The Sexual Revolution? 415
TOM W. SMITH

Comment and Controversy

Comment 436
TOM W. SMITH

Reply 437
JEAN-NOËL KAPFERER

Books

Michael S. Lewis-Beck, *Economics and Elections: The Major Western Democracies* 438
LEE SIGELMAN

Warren E. Miller, *Without Consent: Mass-Elite Linkages in Presidential Politics* 439
BENJAMIN I. PAGE

Samuel P. Oliner and Pearl M. Oliner, *The Altruistic Personality: Rescuers of Jews in Nazi Europe* 442
PAUL D. ALLISON

Andrew M. Greeley, *Religious Change in America* 444
NORVAL D. GLENN

Book Note: Richard G. Niemi, John Mueller, and Tom W. Smith, *Trends in Public Opinion: A Compendium of Survey Data* 447

AAPOR Proceedings

Conference Report 448

AAPOR Award 450

Presidential Address: Polling as Political Science and Polling as Journalism 451
PHILIP MEYER

Conference Program 460

Annual Membership Meeting 473

NOTICE TO CONTRIBUTORS

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1990 issue of POQ.

ATTITUDINAL SOURCES OF PROTEST BEHAVIOR IN FRANCE: DIFFERENCES BETWEEN BEFORE AND AFTER MEASUREMENT

ROY PIERCE AND
PHILIP E. CONVERSE

Abstract Analyses of the attitudinal origins of political protest that rely on surveys conducted *after* the behavior has occurred assume that the attitudes expressed were also present *before* the behavior occurred, and ignore the possibility that postprotest attitudes are the result, rather than the cause, of protest behavior. This article addresses two central questions: (1) How differently may the same protest movement be interpreted, depending on whether the presumed causal attitudes are obtained before or after the protest took place? (2) What kinds of attitudinal transformations are the participants and nonparticipants in protest behavior likely to experience? The data derive from a panel study conducted in France, the first wave of which took place in the spring of 1967, more than a year before the famous May 1968 mass protest occurred, and the second wave of which took place in the summer of 1968, soon after the upheaval, at which time reports of protest behavior were obtained. Identical measures of protest-related attitudes were obtained at each wave. Both the methodological and substantive implications of the differences between "before" and "after" measures of protest-related attitudes are discussed.

There are perils in interpreting the attitudinal origins of protest behavior on the basis of survey data obtained *after* the behavior has oc-

---

ROY PIERCE is Professor of Political Science at the University of Michigan. PHILIP E. CONVERSE is Director of the Center for Advanced Study in the Behavioral Sciences, Palo Alto, California. This is a revised version of a paper originally presented at the 1988 conference of the American Association for Public Opinion Research, Toronto, Canada, May 19–22, 1988. Research for this paper was made possible by a grant (GS-37997) from the National Science Foundation, and by support for data collection from the Ford Foundation, the University of Michigan Council on International Programs, and the Horace H. Rackham School of Graduate Studies of the University of Michigan. We also gratefully acknowledge the assistance of Jean Dotson.

Public Opinion Quarterly Volume 54:295–316 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

curred. Rioters may be distinguished from nonrioters, demonstrators from nondemonstrators, or strikers from nonstrikers but, however plausible the resultant differences may appear to be as explanations for why the protest took place, there always remains the nagging concern that those differences may be the consequences of the protest movement rather than its causes. The rioters, demonstrators, or strikers may be transformed in greater or lesser ways by their participation in the protest movement, just as the bystanders may also have their outlooks altered by their perceptions of the surrounding unrest. Only measures taken *before* protest behavior occurs escape the temporal ambiguity that renders postprotest measures suspect as causal indicators.

That much is obvious. What is less clear is just how different the interpretation of any given protest action, based on postprotest measures, might be from the interpretation of the same events that rested on preprotest measures. Just how much risk do we run of being led astray by postprotest measures or, to put the question more manageably, *what kinds of* risk do we run by doing so? And what kinds of attitudinal transformation are the participants and nonparticipants in protest behavior likely to experience? Those are the issues to which this article is directed.

In an earlier article (Pierce and Converse, 1989) we identified attitudinal roots of the French upheaval of May 1968 distinguishing, on the one hand, between demonstrators and nondemonstrators, and, on the other hand, between strikers and nonstrikers. The analysis rested on data from a panel study and had the rare (if not unique) property of containing attitudinal measures taken on the individual level more than a year before the large-scale protests took place and some of our panel respondents reported joining them. In those circumstances, we were able to demonstrate unambiguously that certain attitudes distinctive to the protesting groups—demonstrators and strikers—were in place well before the May events occurred.[1]

Here we will compare those results, which rested on "before" measures, with those that would have emerged had we been left to rely instead only on "after" measures. This comparison is possible because almost all of the measures obtained more than a year before the 1968 revolt (specifically, in the spring of 1967) were obtained again, after the May events, at the same time as we ascertained whether and in what ways our national sample of respondents had participated in the protest movement. We can, therefore, compare the predictive power of variables that are identical in every respect except that, in one case, they were measured more than a year before the May 1968 events while, in the other case, they were measured shortly after them.

---

[1] The earlier article describes our procedures in detail, reviews the related literature, and discusses our findings in the light of the theory surrounding mass protest.

The Predictor Variables

Four classes of predictor variables were employed for the "before" analysis: conventional political orientations, more diffuse political attitudes, measures of personal satisfaction, and opinions on issues. The measures of political orientation included self-placement on the left-right dimension and attitudes toward Charles de Gaulle. The broader measures of political attitudes included a three-item scale of trust in government, a measure of trust in people, and two measures of political efficacy. The satisfaction items included satisfaction with one's family income, with one's work, and with the adequacy of educational opportunities for young people. The issue opinion items included the fairness of the working class' share of the national income, the protection of union rights, and the adequacy of government efforts on behalf of education. The latter issues were central in the upheaval for worker and student protesters, respectively.

Virtually all of those predictors were retained in identical form for the "after" analysis. An item relating specifically to reform of the university system replaced the earlier item about the adequacy of government efforts in the domain of education generally. The trust in government scale was dropped. Those were the only changes and, as it should become clear, neither of them creates any fatal obstacle to our investigation. (See the appendix for the exact wording of the questions.)

The "Before" Models

The analysis of 1968 French revolt behavior on the basis of attitudes measured in 1967 showed—as conventional accounts at the time suggested—that the mass upheaval consisted of two distinct components: demonstrators and strikers.[2] The demonstrators were comparatively well-educated anti-Gaullist leftists who were also dissatisfied with their jobs and remarkably sensitive to perceived inadequacies in both educational opportunities for young people and government efforts to overcome them. The demonstrations, therefore, pivoted around the two primary elements of French political conflict during the late 1960s: the left-right axis and attitudes toward Charles de Gaulle. The demonstrators' ranks were populated by people who displayed more than a little trace of middle-class radicalism, in the form of their high degree of dissatisfaction with their work (but not with their incomes). And their bond of sympathy with the students who launched the protest movement in the first place lies in the extent to which they perceived social injustice in educational terms.

---

[2] Our procedure for constructing the basic models involved, first, cross-tabulating each variable in our inventory of potential predictors with demonstration participation and strike participation, separately, to determine which were statistically significant on a bivariate basis, using gamma as the test statistic. Then all the statistically significant variables were included in logistic regressions to produce "maximal" models. Finally, the least significant variables were removed successively until we had arrived at optimal

The strikers, by way of contrast, were characterized by quite a different configuration of basic attitudes. Their main point of similarity with the demonstrators was their self-placement on the left-right scale. The strikers were well to the left of the nonstrikers. But the strikers were not particularly anti-Gaullist. They were, however, highly dissatisfied with the working class' share of the national income (although, like the demonstrators, they did not register notable discontent with their personal incomes). They show some traces of discontent with their jobs and educational opportunities, but not to the same extent as the demonstrators. While the demonstrations were an expression of middle-class radicalism closely attuned to education as a focal point of societal conflict, the strikes reflected traditional working-class solidarity.

The protesters of 1968, therefore, were characterized more than a year before the outbreak of the May events by familiar political attitudes. Both demonstrators and strikers were leftists, while the demonstrators were anti-Gaullist as well. Demonstrators and strikers were each distinguished, as groups, by issue positions particular to their group. In a way that almost surely defines their sympathetic ties to the protesting students, the demonstrators were highly sensitive to educational issues. The striking workers, on the other hand, were engaged in a time-honored class action. More diffuse political attitudes, such as trust in government, political efficacy, or trust in people, were not distinctive among either demonstrators or strikers.

---

models in terms of statistical significance and proportions of reductions in predictive error.

For the analysis of demonstration participation, we excluded persons residing in localities where no demonstrations were reported or who lived in the rural fringes of electoral districts where demonstrations were reported in the urban centers, as well as respondents who were more than 50 years of age in 1968.

For the analysis of strike participation, we compared striking workers with nonstriking workers among persons employed by others (the *salariat*).

The "Before" and "After" Models Compared

A very different picture emerges if we apply the same models, but instead of basing them on data obtained in 1967, well before the outbreak of the May 1968 revolt, we base them on the data obtained in 1968, at the same time as our reports of demonstration participation and strike participation were obtained. See Tables 1 and 2 for the two sets of models, relating to demonstration participation and strike participation respectively.[3]

Before commenting on Tables 1 and 2 we should include a terminological note in order to avoid ambiguity. We have already referred to "before" and "after" models. All our models are, of course, "after" models in a strict sense because our dependent variables—demonstration participation and strike participation—necessarily derive from our 1968 survey, which was conducted after the demonstrations and strikes occurred. Our use of the terms "before" and "after" applies only to the time at which the data contained in our independent variables were obtained. Our "before" models are optimized on the basis of attitudinal data obtained in 1967. Our "after" models are optimized on the basis of attitudinal data obtained in 1968. The left-hand panels of Tables 1 and 2 contain our optimized models based on 1967 attitudinal data; the right-hand panels show the results of the same optimization, but based on 1968 attitudinal data.

The richness and distinctiveness of the "before" models evaporate when they are based on 1968 data. For the demonstrations, left-right locations soar in importance, but only a trace of the importance of satisfaction with educational opportunities remains. Attitudes toward de Gaulle, pro and con, almost disappear as factors, and satisfaction with one's work actually becomes a negative, if weak, consideration.[4]

With regard to strike participation, the only factor that retains importance when the "before" model is run with 1968 data is left-right posi-

---

[3] Our original model of demonstration participation, based on 1967 attitudes, included an independent variable referring to opinions about whether the government should do more for education. That item was not repeated in the 1968 survey, so we have omitted it from the model reported in Table 1 to ensure complete comparability. However, a comparison of the original 1967 model with the same model, but based on 1968 data and including an independent variable concerning attitudes toward reforming the universities that was obtained in 1968 but not in 1967, gives virtually the same results as the leaner comparison reported in Table 1.

[4] The variable labeled "Satisfaction with Work Complement" is without intrinsic significance. It was designed to maximize the number of cases that could enter the regression relating to demonstration participation, by allowing us to include both respondents who worked and therefore could register how satisfied they were with their work as well as respondents who did not work (but certainly could demonstrate). One would normally expect the work satisfaction variable and its complement to produce complementary coefficients, as in the case of the "before" model based on 1967 data. We are indebted to Charles Franklin of Washington University for his advice on this matter.

Table 1. Effects of Applying Models Based on 1967 Data with 1968 Data, Model of Demonstration Participation (age 21–50 in Hot Districts)

Predictor Variable	1967 Optimization Based on 1967 Attitudinal Data		1967 Optimization Based on 1968 Attitudinal Data	
	Unstandardized *B*	*B*/SE	Unstandardized *B*	*B*/SE
Left-right location (L)	.454	2.51	.723	3.69
Satisfaction with De Gaulle (−)	.366	2.17	.118	0.65
Satisfaction with educational opportunities (−)	.359	2.45	.258	1.57
Satisfaction with work (−)	.286	1.97	−.114	−0.75
Satisfaction with work complement (+)	−.332	−1.76	−.441	−2.14
R^2 analogue		.158		.172
r (predicted score by actual score)		.46		.49
Chi-sq. estimate (*df*)		36.68 (5)		32.06 (5)
Unweighted *N*		269		263

NOTE: The less trust, satisfaction, or efficacy (−), or the more leftist (L), the greater the probability of engaging in revolt behavior.

Table 2. Effects of Applying Models Based on 1967 Data with 1968 Data, Model of Strike Participation (*Salariat*)

	1967 Optimization Based on 1967 Attitudinal Data		1967 Optimization Based on 1968 Attitudinal Data	
Predictor Variable	Unstandardized *B*	*B*/SE	Unstandardized *B*	*B*/SE
Left-right location (L)	.313	2.65	.291	2.56
Distribution of income unfair to workers (A)	.321	2.12	.046	0.32
Satisfaction with work (−)	.159	1.65	.086	0.84
Government doesn't care (A)	.114	1.41	.064	0.79
R^2 analogue		.050		.026
r (predicted score by actual score)		.30		.22
Chi-sq. estimate (*df*)		15.71 (4)		11.71 (4)
Unweighted *N*		246		241

Note: The less trust, satisfaction, or efficacy (−), or the more agreement with the item (A), or the more leftist (L), the greater the probability of engaging in revolt behavior.

tioning, although it is less strong. The other factors of the 1967 strike model, which give it an intelligible quality, distinct from that of the "before" demonstration model, virtually vanish. The overall explanatory power and goodness of fit of the strike model, which are weak even when it is nourished with 1967 data, fall sharply when 1968 data are employed.[5]

Our efficient models of demonstration and strike participation based on attitudes expressed more than a year before the demonstrations and strikes took place are both substantively satisfying and statistically significant. We have shown that these same models become distorted almost beyond recognition when we base them on attitudes concerning identical subjects, but expressed *after* the protest behavior in question took place.

If, instead of expressing our "before" models with "after" attitudinal data, we follow the same procedure we employed to create the before models (see n. 2), but now use only the after data, we obtain a new set of models. These "after-only" models, shown in Tables 3 and 4, are the models at which we would have arrived if we had had only a single, post–May 1968 survey with which to work. We also present in Tables 3 and 4 what these models, optimized on the basis of 1968 attitudinal data, would look like if we based them, not on those data, but rather on attitudinal data obtained in 1967. This procedure is, of course, just the opposite of that followed in the construction of Tables 1 and 2.

The after models, based wholly on after data, are less distorted substantively when they are injected with the before data, than the before models, based on 1967 attitudinal data, are distorted when the before models are applied with the 1968 attitudinal data, as shown in Tables 1 and 2. However, the overall fit and significance of the hybrid models decline sharply relative to the pure after models.

The main comparison in which we are interested is not between identical models constructed with attitudinal data derived at different times, but rather that between our two pure models, the one that we may call the "before" model displayed in the left panels of Tables 1

---

[5] We would not expect an attitudinal model of strike behavior to be as powerful as one relating to demonstration participation. Joining a demonstration is intrinsically a more personal, individual act than striking. Although our characteristic image of collective behavior is the crowd—the ancestor of the demonstration—the decision whether to join a demonstration can be very much an individual matter, and while the demonstrator's sensations may be strongly affected by mass contagion, there is a real sense in which demonstrating is atomistic behavior. Striking, in contrast, calls into play webs of long-standing relationships, among union members or work groups, so that striking can be much less a matter of individual choice than of collective decision. The difference between strikers and nonstrikers among the *salariat* is not so much personal as organizational; it is less attitudinal than institutional. We assume that wage earners enmeshed in infrastructures oriented toward striking struck, while those outside such frameworks did not strike, even though both groups may have had similar attitudes and opinions.

Table 3. Effects of Applying Models Based on 1968 Data with 1967 Data, Model of Demonstration Participation (age 21–50 in Hot Districts)

Predictor Variable	1968 Optimization Based on 1968 Attitudinal Data		1968 Optimization Based on 1967 Attitudinal Data	
	Unstandardized *B*	*B*/SE	Unstandardized *B*	*B*/SE
Left-right location (L)	.590	3.24	.408	2.45
Politics too complicated (A)	−.261	−2.34	−.007	−0.06
Protect union rights (A)	.414	1.69	.388	1.56
Satisfaction with educational opportunity (−)	.229	1.39	.415	3.06
R^2 analogue		.164		.104
r (predicted score by actual score)		.45		.35
Chi-sq. estimate (*df*)		29.54 (4)		21.20 (4)
Unweighted *N*		271		283

NOTE: The less trust, satisfaction, or efficacy (−), or the more agreement with the item (A), or the more leftist (L), the greater the probability of engaging in revolt behavior.

Table 4. Effects of Applying Models Based on 1968 Data with 1967 Data, Model of Strike Participation (*Salariat*)

	1968 Optimization Based on 1968 Attitudinal Data		1968 Optimization Based on 1967 Attitudinal Data	
Predictor Variable	Unstandardized *B*	*B*/SE	Unstandardized *B*	*B*/SE
Left-right location (L)	.253	2.23	.314	2.76
Politics too complicated (A)	−.160	−2.01	.046	0.57
Protect union rights (A)	.323	2.34	.175	1.32
R^2 analogue		.042		.026
r (predicted score by actual score)		.28		.22
Chi-sq. estimate (*df*)		13.22 (3)		12.19 (3)
Unweighted *N*		243		249

Note: The more agreement with the item (A), or the more leftist (L), the greater the probability of engaging in revolt behavior.

and 2 and the other, which we will call the "after-only" model presented in the left panels of Tables 3 and 4. It is the differences between these two models that are at the center of our concern.

The overall difference between the two models consists of three main elements: the clear distinctions in the "before" analysis between two differently motivated groups of protesters—demonstrators and strikers—is replaced in the "after-only" data by an image of a much more homogeneous protest movement; broad ideological positioning in the form of left-right locations gain in importance while attitudes for and against de Gaulle recede as apparent causal factors; and political efficacy, which was totally absent from the earlier models, emerges as a powerful force across the board.

In the before models, the demonstrators' great dissatisfaction with their work and their preoccupation with educational issues clearly distinguishes them from the strikers, whose main policy concern is the distribution of income. This difference almost wholly evaporates in the after-only models. Instead, the issue of union rights, virtually absent from the before models, becomes a powerful factor common to the after models for each protesting group.

At the same time as this shift in policy factors, there is a change in the role of the purely political variables. Attitudes for and against de Gaulle play an important part in the before model for the demonstrators, but Gaullism (and its opposite) disappear from the after-only model. Left-right positioning, in contrast, increases in importance for demonstrators and at least holds its own among the strikers.

Lastly, and in some ways most remarkably, political efficacy was an after factor of unquestionable standing for both demonstrators and strikers. None of our before measures of political trust, personal trust, or political efficacy had registered more than marginally as a forward predictor of protest behavior. In the after measurements, however, our single indicator of political efficacy proved to be a potent factor indeed. When presented with the statement: "Sometimes politics and government seem so complicated that a person like me can't really understand what's going on," the demonstrators and strikers were much more likely to disagree than were their passive fellow citizens. The protesters, who had not notably differed from the rest of the population in this regard in the spring of 1967, had little difficulty figuring out what was going on politically in the summer of 1968.

A Closer Look at Attitudinal Shifts

The different models that we have outlined, contrasting the relationships between various attitudinal attributes and behavior during the

French upheaval of May 1968, depending upon whether those attitudes were measured in 1967, more than a year before the May events, or shortly after those events, obviously reflect substantial attitude changes between those two moments.[6] We can gain a better understanding of the nature of those attitudinal shifts by examining the distributions along the relevant dimensions of the protesters, on the one hand, and the passive bystanders, on the other hand.

We will employ two related, but different, measures for this purpose. The first is eta, which expresses the amount of variance *between* the two groups (participants and nonparticipants) divided by the sum of that variance plus the amount of variance *within* each of the two groups. The measure is, in short, a convenient indicator of the degree of polarization between the two groups, and has the useful property of ranging from zero to one. By comparing the etas for the same variable in 1967 and 1968, we can note at a glance whether there was an increase or a decrease in attitudinal polarization between protesters and bystanders.

The second measure consists of the standardized scores of the same two groups on each relevant variable. While the etas give us precise measures of changes in polarization between 1967 and 1968, the standardized scores for the two years indicate the relative contribution of each group to whatever change occurred. These measures are set forth in Table 5.

The most striking phenomenon is the heightened polarization along the left-right dimension between demonstrators and nondemonstrators. The data relating to demonstrations that are presented in Table 5 derive from our subset of districts where there was demonstration activity, including only respondents who were under 50 years of age (the great majority of the protesting demonstrators), but the increase in left-right polarization was a nationwide phenomenon, if less marked between the demonstrators and the rest of the adult population than within our subsample.

The fact that such polarization increased more sharply among the

---

[6] Since we have measurements at only two times, we cannot of course be sure of the exact timing of some of the attitudinal shifts that our data assure us took place differentially between protesters and bystanders over this span of time. It is always possible that the change occurred earlier than we think, i.e., that measurements taken in March 1968, just before the revolt but while it was still unimaginable, might have looked in these regards more like our "after" measurements in the summer of 1968 than like our true "before" measurements drawn a full year earlier. This seems implausible on the face of it, since there is no other obvious cause for these changes than the revolt; and as we are about to see, most of the changes are in fact highly plausible as effects of revolt observation and/or participation. Moreover, in the one important case of left-right polarization, other monitoring data exist that show the brusque polarization changes only after the revolt is well along. Nevertheless, we cannot in all instances be sure that we are dealing with effects of the revolt.

Table 5. Polarization Between Protesters and Bystanders on Selected Predictor Variables, According to "Before" and "After" Measures

Predictor Variable		Eta	Standard Scores		Eta	Standard Scores	
			Demonstrators[a]	Bystanders		Strikers[b]	Nonstrikers
Left-right location	67	.202	−.53	+.08	.201	−.24	+.17
Left-right location	68	.313	−.85	+.12	.204	−.25	+.17
Protect union rights	67	.164	+.43	−.06	.130	+.16	−.11
Protect union rights	68	.191	+.50	−.07	.204	+.25	−.17
Politics too complicated	67	.012	−.03	+.00	.038	+.05	−.03
Politics too complicated	68	.159	−.42	+.06	.099	−.12	+.08
Distr. of income unfair	67	.119	+.32	−.04	.191	+.23	−.16
Distr. of income unfair	68	.109	+.29	−.04	.098	+.12	−.08
Satis. educ. opp.	67	.212	−.56	+.08	.120	−.14	+.10
Satis. educ. opp.	68	.134	−.35	+.05	.118	−.14	+.10
Do more for education	67	.156	+.41	−.06	.073	+.09	−.06
Reform the universities	68	.064	+.17	−.02	.094	+.11	−.08
Satisfied with work	67	.154	−.36	+.07	.140	−.17	+.12
Satisfied with work	68	.008	+.02	−.00	.092	−.11	+.08
Attitude toward De Gaulle	67	.259	−.68	+.10	.151	−.18	+.12
Attitude toward De Gaulle	68	.237	−.63	+.09	.176	−.21	+.14

[a] Among respondents aged 21–49 in 1967 residing in "hot" areas.
[b] Among respondents in the work force employed by others in 1968.

young in the "hot" districts than nationwide is an indication that the shift was not something that occurred prior to the upheaval, at some point between our 1967 observations and the outbreak of the May events, but was rather the result of the disorder itself. Elsewhere (Converse and Pierce, 1986: chap. 15), we have charted the shifts in subjective left-right locations within the French electorate between 1967 and 1968 as a function of proximity to the demonstrations and strikes. Here we have confirmatory evidence that self-perceptions of left-right placement polarized more acutely at the main scenes of the action than elsewhere.

While the leftward movement of the demonstrators is the most impressive figure presented in Table 5 in absolute terms, in relative terms, the shift to the right of the nondemonstrators is almost equally noteworthy. The demonstrators were driven leftward by the experience of protest behavior, while the nonparticipants were, on the average, pulled to the right by their uncertainties and fears, a backlash to the protest events that registered as well in the unprecedented tidal wave of support the Gaullists received at the polls in the wake of the disturbances. The magnitude of the rightward move of the nondemonstrators, which is expressed in Table 5 in standardized form, is all the more striking as the bystanders outnumber the demonstrators in our subset of hot districts by seven to one, and therefore contribute overwhelmingly to defining the mean on the basis of which the scores are standardized. Hence, any significant move away from zero among the nondemonstrators reflects an unusual disturbance of the normally inertial distribution of the electorate along the left-right axis.[7]

The second attitudinal shift that deserves attention is the increase in polarization between 1967 and 1968 over the issue of union rights, whether one considers demonstrators and nondemonstrators, or strikers and nonstrikers. Attitudes toward unions, interacting with the others that we tested, became clearly associated with revolt behavior when they had not been before, both in the case of the demonstrations and in that of the strikes.

Here, too, we almost certainly have an effect of the upheaval of May 1968, as opposed to one of its causes. Apart from student organizations and various small, dogmatic leftist groups that the French refer to as *groupuscules*, with which the French mass electorate was almost completely unfamiliar, the only organized groups that were constantly before the public eye during the May upheaval were the opposition unions, whose leaders issued almost daily pronouncements. Except for

---

7. There is independent evidence, in the form of surveys conducted at the time by IFOP (*Institut Français d'Opinion Publique*), of a very rapid shift in Parisian opinion during May 1968 from sympathy for the demonstrating students to opposition toward them. See *Sondages*, 1968:75–83.

the Communist Party and the small *Parti socialiste unifié* (PSU), the latter also being invisible to much of the population, the political parties were not conspicuous for public efforts to affect the course of events. The unions, therefore, particularly the Communist-led *Confédération Générale du Travail* (CGT), the Catholic-led, "new leftist" *Confédération Française Démocratique du Travail* (CFDT)—which was most outspokenly in favor of the demonstrating students—and the several teachers unions, all of which were opposed to the Gaullist regime, were highly salient organizations in terms of which people could easily orient themselves toward the events unfolding in the streets.

By the time we made our postupheaval observations, virtually the entire French population had positioned itself with reference to the unions, as a function of the behavior and pronouncements of those unions during the upheaval itself. Demonstrators and strikers were well above the mean in the level of their support for the unions; nondemonstrators and nonstrikers were below the mean in that regard. In the circumstances, however, it would be a post hoc fallacy of the most flagrant kind to affirm that those attitudinal differences were prime causal factors in the May upheaval. The demonstrators and strikers were, after the May events, much more favorable to unions than the bystanders were, but these different perspectives were more the result than the cause of those events.

The most remarkable manifestation of the way in which the May events transformed people's attitudes appears in connection with our indicator of political efficacy: expressions of agreement or disagreement with the statement that "politics and government sometimes seem so complicated that people like me can't really understand what is going on." In the spring of 1967, reactions to that assertion were almost randomly distributed between the future protesters, on the one hand, and the future bystanders, on the other.

By the spring of 1968, however, shortly after the May events, there is marked polarization over that issue between demonstrators and nondemonstrators and between strikers and nonstrikers. In fact, when the strikers are contrasted with the nonstrikers, there was actually a reversal of signs between 1967 and 1968. In the earlier year, the future strikers felt more at sea with regard to politics and government than the future nonstrikers did, while in the later year—after the upheaval—the strikers thought that they understood government and politics better than nonstrikers did, and by a fairly wide margin.

One would be hard put to account for these dramatic transformations without reference to the force unleashed by the May upheaval itself. Back in 1967, the future protesters might not have had much confidence in their ability to untangle some of the complex strands of

French politics, but, in 1968, a few days of demonstrating (and possibly also a few nights of rioting) or a few weeks on the picket line went a long way toward clarifying the situation for them. Before the fact, the protesters might well have felt bewildered by the arcane language and ceaseless maneuvers of their political leaders in 1967, but in the spring of 1968 they would have had to be unusually obtuse not to have known "what is going on." Demonstrations and strikes were going on, and they themselves were causing them. The protesters of 1968 gained confidence and a sense of political efficacy from the protest itself, even if they failed in most respects to achieve their goals.[8]

What were these goals? It is virtually impossible to answer that question on the basis of our *postrevolt* data. The policy matters that clearly distinguish demonstrators from nondemonstrators, strikers from nonstrikers, and demonstrators from strikers, when we employ the *prerevolt* data almost completely drop from sight after the upheaval. In the "before" measures, the strikers differ from the nonstrikers by their greater conviction that the distribution of income is unjust, but in the "after-only" measures, the degree of polarization over that issue has greatly receded. This is not because the nonstrikers moved closer to the position of the strikers on the issue between 1967 and 1968, while the strikers remained adamant about the distributive question. The nonstrikers did move toward the mean, but so did the strikers. It is as though the issue had lost importance for them even as their efforts reached the highest point of intensity.

---

[8] We do not mean to suggest that the opinion changes between 1967 and 1968 were entirely the result of the May upheaval. In order to measure how much of our respondents' 1968 attitudes could be accounted for by their attitudes on the same subject in 1967, their demonstration and/or strike participation, and the 1967 variables that we found to be related to demonstration and strike participation, we ran two sets of multiple regressions with those items as the independent variables. One set of regressions included respondents under 51 years of age and residing in hot districts; the other consisted of persons employed by others (whether in the private or public sector). These two sets of respondents match the two groups that appear repetitively in Tables 1 through 5, relating to demonstration and strike participation respectively. For the people under 50 in hot districts we included demonstration participation as a dummy variable, but not strike participation, because the set includes many nonworkers. For the employed, however, we included both demonstration participation and strike participation as dummy variables, as workers could demonstrate as well as strike. For each set, the dependent variables were left-right location in 1968, attitude toward unions in 1968, and political efficacy in 1968 (as measured by the "politics too complicated" variable). Those are the three items on which opinion changed most dramatically between 1967 and 1968.

Space limitations prevent us from reproducing the six regressions. We simply summarize them. The total variance accounted for ranged from 22 percent to 45 percent. In every case, the independent variable most closely associated with the dependent 1968 attitude is the attitude on the same subject in 1967. Demonstration participation is significantly related to the dependent variable in each case except that concerning unions among persons aged 21–50 in hot districts. This is not surprising; we have argued that attitudes toward unions were shaped by the unions' pronouncements and behavior more than by the behavior of our respondents. Among the *salariat*, however, demonstration participation was a significant factor in all three regressions, and strike participation was significantly related to political efficacy and to attitudes toward unions, but not to left-right locations. Attitudes that we found in our "before" models to be related to demonstration and strike participation were sometimes also related to 1968 attitudes, but more often they were not. Copies of the complete set of regressions will be furnished on request.

Similarly, our 1967 data show that the demonstrators were characterized by their dissatisfaction with educational opportunities in France and with government efforts in the educational domain. Only a modest trace of the demonstrators' sensitivity to education as a litmus test of social justice appears in the postrevolt data. Even the 1967 indications that the demonstrators were unusually dissatisfied with their work, consistent with "new working class" themes, disappeared with the occurrence of the revolt. Indeed, here again we encounter a change in signs; after the upheaval the demonstrators were slightly *more* satisfied with their work than the nondemonstrators.

The final, but not the least interesting, phenomenon that requires comment is the changed role of opinions about Charles de Gaulle as a factor in the revolt, depending upon the moment of observation. In 1967, among young adults in the districts where there was demonstration activity, polarization of attitudes toward de Gaulle was high and anti-Gaullist attitudes were a source of demonstration participation. Anti-Gaullism was a lesser factor in the strike movement.

At the time of our second wave of interviews, after the May events, there was discernibly greater polarization of mass attitudes pro- and con de Gaulle across the total electorate, but there was a modest *decrease* in such polarization between demonstrators and bystanders within our subset of young people in hot districts. This slight dip in polarization, *when considered in an interactive sense with* our *other* potential predictor *variables,* causes anti-Gaullist attitudes actually to drop out of the equation as a presumed causal element in the demonstrations. This is astonishing, given the extent to which French political conflict revolved around attitudes toward de Gaulle during the late 1960s, and especially during the 1968 crisis. It illustrates perhaps better than any other of our findings how misleading an analysis based on postprotest data alone could have been.

Among the people under 50 in the districts where demonstrations were later to occur, the battle lines between the pro–de Gaulle and the anti–de Gaulle forces were clearly drawn more than a year before the historic outburst occurred. Indeed, it is likely that the intensity of the attitudinal conflict pivoting around de Gaulle was already near fever pitch at that time. When the May events erupted, attitudes toward de Gaulle among those people did not polarize further. Instead, and here our detailed discussion of the individual predictor variables comes full circle, those attitudes were transposed onto the left-right dimension, whose elasticity was patently greater than that of the Gaullism dimension which it surpassed in salience. It was along the left-right dimension—not the Gaullism dimension—that polarization increased between demonstrators and onlookers. If we rely on postrevolt data, we come away believing that the demonstrations were essentially a leftist revolt in which anti-Gaullist attitudes played no significant role. Such an interpretation would be absurd in the light of the most elementary familiarity with postwar French politics generally, or the May events in particular. Fortunately, we have more than elementary information on which to rely. We have observations made over a year before the May upheaval, and those are consonant with what both expert opinion and common sense suggest occurred in France in 1968.

But why did left-right self-perceptions crowd out attitudes toward de Gaulle in our postrevolt measures? Conventional wisdom would suggest that it should have been the other way around. One would normally expect that in moments of acute social and political crisis the mass public would be more responsive to personal leadership, pro or con, than to abstract symbols. But such a reading would overlook the extent to which left-right imagery had penetrated the public mind in France. The greater left-right polarization in 1968, compared to 1967, does not so much represent attitudinal change as it reflects the mobilization of opposing political forces that took place during the May upheaval and the subsequent electoral competition in late June and July. The heightened left-right polarization was more than simply one factor among the several we have examined; it was, in a sense, the summary of them all. In the short space of time between the May outburst and the July truce, practically all the elements of the left-right repertory appeared on the political stage. The confrontation of May and the elections that followed during the summer evoked layers of political memory, of which left-right self-perceptions represent the simplest, most universal expression. People were virtually compelled to choose sides; and when they did, they did so in familiar left-right terms.[9]

Taken together, the shifts in attitudes between our first moment of observation, more than a year before the May upheaval, and the second, after the May events had transpired, suggest a kind of homogenization of what, on the basis of the "before" analysis alone (and apparent even to casual observation at the time of the events), amounted to two separate movements. The demonstrations were the work of comparatively well-educated, anti-Gaullist leftists, who were unsatisfied with their jobs and particularly concerned with education, either in its own right, or as a symbol of the state of justice in French society. The strikes were an expression of left-wing, working-class solidarity, oriented primarily to the issue of distributive justice. All that emerges clearly from an examination of the "before" data.

---

[9] Strictly left-right competition became the dominant form of political conflict in postwar France during the mid-1960s. It was expressed most clearly at the presidential election of 1965, when François Mitterrand was the single candidate of leftist parties united in opposition to de Gaulle. Left-right conflict also characterized the legislative elections of 1967, when two-thirds of the second-ballot contests were straight fights between a rightist and the district's "single leftist candidate." The unity of the left-wing parties suffered strain during the May upheaval, but they maintained their electoral alliance for the 1968 legislative elections, even though it failed badly as the result of a rightist backlash to the May demonstrations.

When we look at the "after" data, however, that sharp distinction between two different, discontented groups of French citizens, motivated by separate considerations, blurs and virtually disappears. On the basis of the "before" measures, the only factor that is unequivocally common to the two groups of protesters—demonstrators and strikers—is the left-right dimension; the other main attitudes underlying the upheaval tend to cluster around one group of protesters or another. When we examine the "after" measures, we find that left-right locations are still common to the two groups (in much stronger form for the demonstrators, as we have seen), but that attitudes toward union rights and a sense of political efficacy (totally absent from the significant "before" factors) are also associated with both demonstrating and striking.[10] The only indications that the demonstrations and strikes were not cut from identical cloth are the weak expression, among the demonstrators, of concern over the lack of educational opportunity for young people in France.

We should not, however, take this blurring of the two streams of protest when "after" measures are used to discriminate protesters from bystanders, to mean that in some more general way we get a good level of prediction only when using genuinely prior attitudinal variables, and little or no "prediction" with the "after" attitudes. In fact, it is an intriguing property of these data that our best indicators of overall predictability, the R^2 analogue and the correlation between actual and predicted scores on the dependent variable given in the panels of Tables 1 and 2 and Tables 3 and 4, show what is in almost uncanny degree about the same level of predictability whether we use the "before" or the "after-only" model. This is surprising mainly because the configuration of useful predictors is so different in the two cases that some overall discrepancy in predictability might be expected. Instead, we find a noteworthy parity. This is a somewhat insidious circumstance: we would not be tipped off that the "after" model was as inadequate as we suppose it to be by the presence of feeble levels of "prediction" using only after measures. It "accounts" as well for variance in protest behavior as the before model, and hence is actively misleading.[11]

---

[10] Finkel (1987) found that among West Germans, peaceful protest had no effect and "aggressive" behavior had a negative effect on political efficacy. However, the data he analyzed differ from our own on every dimension—sample frame, behavioral indicators, and indicator of political efficacy—so no direct comparison may be made.

Conclusions

Our examination of two contrasting forms of analysis of political protest, one based on attitudinal measures taken well before the protest in question, the other resting on attitudinal measures taken at the same time as the protest behavior itself was ascertained, leads to two kinds of conclusions: one relating to method, the other concerning substance. Both have daunting implications for research on the attitudinal sources of mass behavior in extreme situations.

With respect to method, it appears beyond doubt that "after" measurement is seriously contaminated by the experience of the protest itself. Protesters and bystanders alike are transformed by the experience of protest to the extent that postprotest measures are highly unreliable as indicators of the underlying causes of the protest behavior. Our "before" measures supply a much more complete and plausible account of the attitudinal origins of the French upheaval of May 1968.

We do not overlook the fact that the generation and evolution of mass opinion is a continuous process. We have frozen that process at two moments, with the results that we have reported. We might have chosen dozens of other times at which to question our respondents, and we have no way of knowing what the results would have been. But it seems reasonable to hold that as between some 13 months *before* the events, and some 2 or 3 months *after* the events, the earlier point is preferable. We would not argue that it would have been better still to have made our initial observations 48 months before May 1968, and it might well be that the ideal time to have probed the French popular political psyche was April 1968, rather than some longer time before the historic outburst occurred. But if our moment of "before" observation was not ideal, neither was it ill-chosen, while we have good reason to believe that *any* "after" measure is fraught with misleading error.

Of course, researchers may not be particularly interested in the causes of protest (as we are here). They may be concerned instead primarily with its consequences, and among these the opinions and behaviors of people who have experienced it in one form or another. In this case, postprotest measures are quite appropriate, and if they can be linked with still later measures of opinions or behavior, they may be treated as causes themselves.

---

[11] On the basis of our chi-square estimates, the "after-only" model for demonstration participation is even more significant statistically than the "before" model.

As for the substance of the kinds of transformations that widespread protest produces, our data indicate that mass protest does wonders to concentrate the mind. The hoariest political symbols take on renewed importance. Mass opinion crystallizes around the most visible and vocal groups that are the fulcrums of political action. The activists see the situation clearly and have no doubts about the rightness of their cause. But the issues over which the protesters initially took to the streets recede in importance. Like the Grecian warriors at Troy, the combatants forget why they took up arms in the first place.

Appendix

QUESTION WORDING

Trust in Government Scale (1967 Only)

Croyez-vous que les gens qui nous gouvernent donnent une chance égale à tous, qu'il s'agisse de gros bonnets ou de gens ordinaires, ou croyez-vous qu'ils accordent plus d'attention aux gros bonnets?

Croyez-vous que les gens qui nous gouvernent gaspillent l'argent de nos impôts (une grande partie, une certaine partie, ou pas beaucoup)?

Jusqu'à quel point, selon vous, peut-on faire confiance aux gens qui nous gouvernent pour agir comme il faut (à peu pres toujours, la plupart du temps, quelquefois, ou pratiquement jamais)?

Trust in People

D'une façon générale, diriez-vous qu'on peut faire confiance aux gens, ou au contraire qu'on ne prend jamais assez de précautions avec certaines personnes?

Political Efficacy

Je ne crois pas que le gouvernement se préoccupe beaucoup de ce que les gens comme moi pensent. (D'accord, pas d'accord).

La politique et le gouvernement semblent parfois si compliqués que les gens comme moi ne peuvent vraiment pas comprendre ce qui se passe. (D'accord, pas d'accord).

Satisfaction

Diriez-vous que, dans l'ensemble, vous êtes très satisfait, assez satisfait, peu satisfait, ou pas du tout satisfait:

Du revenu de votre foyer?

Du travail que vous faites (ou, si vous ne travaillez pas, du travail que fait le chef de famille)?

Des possibilités qu'ont les enfants de nos jours de recevoir toute l'instruction dont ils ont besoin?

Opinions on Issues

Voudriez-vous nous donner votre opinion, en disant si vous vous estimez d'accord ou non? Et, comme avant, si vous n'avez pas d'opinion, cela nous intéresse aussi.

Income Distribution

Dans la répartition du revenu national les ouvriers sont vraiment défavorisés.

Education

L'Etat devrait faire un plus gros effort pour l'enseignement. (1967 Only)

L'Etat devrait réformer l'université. (1968 Only)

Union Rights

Les droits des syndicats devraient être mieux défendus.

References

Converse, Philip E., and Roy Pierce (1986)
Political Representation in France. Cambridge: Harvard University Press.
Finkel, Steven E. (1987)
"The effects of participation on political efficacy and political support: Evidence from a West German panel." The Journal of Politics 49:441–464.
Pierce, Roy, and Philip E. Converse (1989)
"Attitudinal roots of popular protest: The French upheaval of May 1968." The International Journal of Public Opinion Research 1:221–241.
Sondages: Revue française de l'opinion publique (1968)
No. 2, pp. 75–83.

THE PRESIDENTIAL EFFECT: THE PUBLIC HEALTH RESPONSE TO MEDIA COVERAGE ABOUT RONALD REAGAN'S COLON CANCER EPISODE

MARTIN L. BROWN AND
ARNOLD L. POTOSKY

Abstract Little previous research has been done on the public health impact of mass media coverage of cancer episodes of public figures. This paper uses a variety of data sources to examine the impact of President Reagan's colon cancer episode of July, 1985. Records of phone calls to the Cancer Information Service of the National Cancer Institute are examined as a measure of public interest and concern about colorectal cancer; data on the use of two colorectal early detection tests—proctoscopy and fecal occult blood tests—are looked at as a measure of behavioral change; and data on the incidence of early and advanced colorectal cancer are used to estimate the potential public health impact of this behavioral change. We find that there was a sharp, albeit somewhat transitory, increase in public interest in colorectal cancer in the wake of President Reagan's colon cancer episode, with a corresponding increase in the use of early detection tests. The incidence data on early and advanced disease is indicative of a beneficial public health impact, but this can be confirmed only after additional data on mortality becomes available.

Introduction

"Reagan Tumor Removed; No Cancer Evident," read the banner headline of the *Washington Post* of July 14, 1985. The front page article detailed the nature of the surgical procedure, the removal of a two-foot long segment of the President's colon and went on to discuss the nature

MARTIN BROWN is Economist and ARNOLD POTOSKY is Operations Research Analyst, both at Applied Research Branch, Surveillance Program, Division of Cancer Prevention and Control, National Cancer Institute. The authors wish to thank Larry Kessler and Edward Sondik for helpful comments and Michael Anderson for extracting data from CIS records.

Public Opinion Quarterly Volume 54:317–329 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

of his condition, "a villous adenoma, the type of intestinal tumor most likely to be cancerous" (Rensberger, 1985). Across the nation, this story was replicated in the print and broadcast media, with accompanying background pieces on the nature of colon cancer and on methods of early detection, diagnosis, and treatment (for example, Stoler, 1985).

The purpose of this paper is to explore the public response to this intense dissemination of information about the early detection and treatment of colon cancer and assess the possible impact on the public's health. The response to this unplanned and unanticipated informational campaign may shed some light on the expected response to intentional public information campaigns about cancer prevention and early detection, or on the expected response if the National Cancer Institute or the American Cancer Society attempts to take advantage of similar chance events that may occur in the future. Our analysis is based on the examination of several types of nonreactive measures. This a departure from the more usual approach of using attitude surveys to analyze media impact.

Background

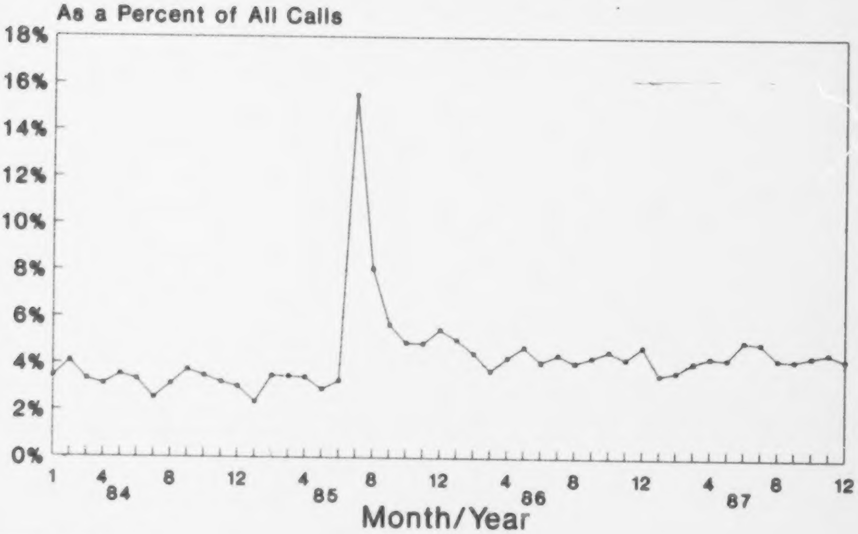
Surprisingly little is known about the effect of mass media campaigns on the knowledge, attitudes, and behavior of the public toward preventive health. As Warner puts it, "The literature on the subject is thin and decidedly inconclusive" (Warner, 1987). Flay (1987) has reviewed the literature on the evaluation of 40 mass media programs/campaigns designed to influence cigarette smoking. While acknowledging that the evaluation methods used were of mixed quality, he concludes that the necessary components of a successful mass media public health campaign are: (1) novelty and diversity of spot announcements, (2) widespread dissemination, (3) high saturation, and (4) endurance. The mass media publicity surrounding President Reagan's colon cancer met only the second and third of these criteria. The extremely high level of the dissemination and saturation and the dramatic nature of the event itself, however, may have compensated for the lack of the other conditions.

There are several studies that have evaluated the effects of mass media campaigns on cancer prevention or screening behavior. Levy and Stokes (1987) evaluated the effects of a national advertising campaign by Kellogg Company promoting its high-fiber cereal. The campaign included a National Cancer Institute statement about the role of fiber in cancer prevention. Levy and Stokes found that the sales of all high- and moderate-fiber cereals increased after the campaign, while the sales of many low-fiber adult brand cereals (notably granola-type cereals) decreased. Doherty and MacKie (1988) have studied the effects of a mass media campaign in Scotland, designed to encourage the early detection of melanoma, the most lethal type of skin cancer. They were able to document an increase in the number of people seeking advice from primary care practitioners and being referred to dermatological specialists for suspected skin cancer. More importantly, they found an increase in the total number of melanomas diagnosed and an increase in the proportion of early stage, treatable melanomas detected. Fink et al. (1978) reported a short-lived increase of participation in the HIP breast cancer screening study following the national media coverage of Betty Ford's and Happy Rockefeller's breast cancer disease in 1973. Heath et al. (1989) looked at monthly data from the Behavioral Risk Factor Surveillance System on the utilization of screening mammography for the early detection of breast cancer just prior to, and after, the diagnosis of Mrs. Reagan's episode of breast cancer in October, 1987. For the months of January through October, the usage rate of mammography averaged about 28%. For November and December this rate increased to 38%. This is a suggestive result, although the data are very limited and further study of this event is warranted.

Methods

This paper is novel because, in order to gauge the effect of the media blitz surrounding President Reagan's colon cancer episode, we examine three independent types of information. The first, phone calls from the general public to the Cancer Information Service (CIS) of the National Cancer Institute (NCI), reflects the degree of public interest aroused about colon cancer. The second, rates of utilization of the two main early detection techniques for colon cancer—proctoscopy and fecal occult blood tests (FOBT)—measures behavioral change resulting from information assimilation about early detection of colon cancer. Lastly, the incidence of early and advanced colorectal cancer will be examined prior to and after July, 1985, using data from NCI's Surveillance, Epidemiology, and End Results (SEER) program. An increase in the use of early detection tests should result in a measured increase in overall incidence of the disease, an even larger increase in the incidence of early stage disease and possibly, with some time delay, a reduction in the incidence of late stage disease. This shift in incidence of early and advanced disease provides an estimate of the potential public health benefit resulting from the change in public attitudes and behavior.

This methodology, of using several types of nonreactive data and



SOURCE: CIS, NCI

Figure 1. Calls To Cancer Information Service

comparing them for cross-validation, is a departure from the more conventional use of attitude surveys to gauge media impact. While our method has some pitfalls, as we note, it has the merit of being based on data that reflect behavior as well as intent, and avoids some of the biases of attitude survey data, such as the framing effect.

Results

CALLS TO THE CANCER INFORMATION SERVICE

As part of its Office of Cancer Communications, the National Cancer Institute maintains the Cancer Information Service, a nationwide network of regional information resource centers that can be reached by the public through the toll-free number 1–800–4–CANCER. A standardized form in use in all CIS offices makes it possible for CIS to keep a record of the 30,000 to 40,000 phone calls received monthly, the nature of each inquiry, and the characteristics of the caller.[1] Figure 1 shows the proportion of CIS calls from new callers which were concerned with colon or rectal cancer over the period January, 1984, to

---

[1] All calls are recorded. Detailed demographic and other data are collected on a sample of calls at each CIS office. The data reported here exclude repeat callers.

December, 1987.[2] The most striking feature of Figure 1 is the sharp increase in inquiries related to colon and rectal cancer in July, 1985, coinciding with the media attention to President Reagan's colon cancer. The percentage of calls about colon cancer increased more than four-fold from June to July, 1985. While much of this interest subsided within the next three months, it is notable that the post-July, 1985, level of calls about colon cancer remained permanently elevated at slightly above the pre-July, 1985, levels for the entire period shown in Figure 1. For example, the average monthly number of calls for colorectal cancer for the period January, 1986, to December, 1987, was 1356 calls, as compared to 1013 calls per month for the period January, 1984, to June, 1985. The media attention to President Reagan's colon cancer appears to have had a long-term as well as a short-term effect on the public's interest in this disease.[3] The long-term persistence in the increased interest in colorectal cancer is notable when it is considered that in the post-July, 1985, period, there was competing publicity about other dramatic developments in the cancer area, such as publicity, in December, 1985, about experimental treatment with Interleukin-II (DeVita, 1986).

UTILIZATION RATES OF EARLY DETECTION TESTS FOR COLORECTAL CANCER

Three clinical tests have been recommended by NCI and the American Cancer Society for the early detection of colorectal cancer, digital rectal examination, fecal occult blood tests (FOBT), and proctoscopy.[4] The digital rectal examination is recommended by NCI and ACS primarily as an early detection test for prostate cancer, but it is capable of detecting a small proportion of early colorectal cancers. Unfortunately, data on the utilization of these tests by the eligible populations just prior to and after July, 1985, are not available.[5] Data on utilization

---

[2] In order to obtain a reliable pre-event baseline estimate, CIS calls were actually tracked back to January, 1983. Analysis of this data showed no trend prior to July, 1985. For ease of presentation the data for 1983 and 1984 are not shown in Figure 1.

[3] To put the short-term and long-term effects in perspective the following comparison is useful. The "excess" number of calls over the pre-July, 1985 level was about 1950 calls per month for the months July–October, 1985. For the months November, 1985, to December, 1987, there were about 350 "excess" calls per month.

[4] A fecal occult blood test is a chemical test for detecting blood in the stool. It is usually administered through a take-home, self-test kit. A proctoscopic examination is an examination of the colon usually using a flexible fiber-optic instrument.

[5] Data from American Cancer Society–Gallup Surveys do show an increased use of FOBT between 1983 and 1987. In 1983 about 28% of adults age 50 and over reported ever having had an FOBT; in 1987 this proportion was about 45%. There is no way to know, however, whether this increase was a result of the Presidential Effect or reflects a secular trend due to other factors.

Table 1. Medicare Utilization Rates for Proctoscopy (Procedures per 1,000 Medicare Beneficiaries)

Procedure	1983[a]	1984[a]	1985[a]	1986[a]	1987[b]
Colonoscopy	12.02	15.11	21.32	26.51	21.30
(Diagnostic only)					10.90
Sigmoidoscopy	3.25	4.59	8.01	10.27	25.90
(Diagnostic only)					22.30
Proctosigmoidoscopy	19.45	17.78	18.47	14.28	15.40
(Diagnostic only)					14.20
Total	34.72	37.48	47.80	51.06	62.60
(Diagnostic only)					47.40
Increase over preceding year	—	7.95%	27.53%	6.82%	22.60%

[a] Medicare Part B claims for Alabama, Connecticut, Washington, and Wisconsin (Mitchell et al., 1988).
[b] BMAD Procedure File 1987, p. 105 (provided by HCFA).

rates for proctoscopy under Medicare, however, are suggestive. These data are summarized in Table 1. Table 1 shows utilization rates for diagnostic colonoscopy (full examination of the colon), sigmoidoscopy (examination of between 20 and 60 cm of the colon), and proctosigmoidoscopy (less than 20 cm of the colon). Only the latter two procedures would qualify as primary early detection procedures, i.e., procedures used to screen asymptomatic individuals. It makes sense, however, to consider the combined utilization rates of these three procedures for several reasons. Colonoscopy is often a follow-up procedure to a positive sigmoidoscopy or proctosigmoidoscopy as well as to positive digital rectal or FOBT examinations. Thus, an increase in colonoscopy may reflect an increase in these other procedures. There is also evidence that, until 1987, many physicians were reporting sigmoidoscopies to Medicare as colonoscopies.[6] For these reasons we will look at the combined utilization rates for all these procedures. Table 1 shows the percentage increase per year in these rates. Over the

---

6. This "up-coding" is reflected in the data in Table 1. In 1987 the Health Care Financing Administration implemented coding changes that made it more difficult to misreport sigmoidoscopies in this manner. Notice that between 1986 and 1987 the reported utilization rate for colonoscopies declined while the utilization rate for sigmoidoscopies increased sharply. Since colonoscopies are reimbursed at about $400 while sigmoidoscopies are reimbursed at about $125, this represents a considerable savings to the Medicare budget.

entire period 1983–1987 these procedures grew, on average, at about 12% per year. This pattern was not uniform over time. The largest annual increase, about 28%, occurred during the same year in which President Reagan's colon cancer episode occurred, 1985.[7]

The most widely used FOBT product, with approximately 80% of the market, is Hemmocult produced by Smith-Kline. National sales of Hemmocult to the hospital/private laboratory market were significantly elevated in 1985 compared to 1984 and 1986 (personal communication, Smith-Kline product manager).

INCIDENCE OF EARLY AND ADVANCED COLORECTAL CANCER

The purpose of early detection tests for cancer is to diagnose the disease at an earlier, and apparently more treatable stage. In order for the attendant increase in public interest and preventive behavior to have any direct public health implications it is necessary to see a relative increase in early disease and, later, a relative decrease in advanced disease.[8] It may be useful, therefore, to examine the incidence of early and advanced colorectal cancer prior to and after President Reagan's colon cancer episode.

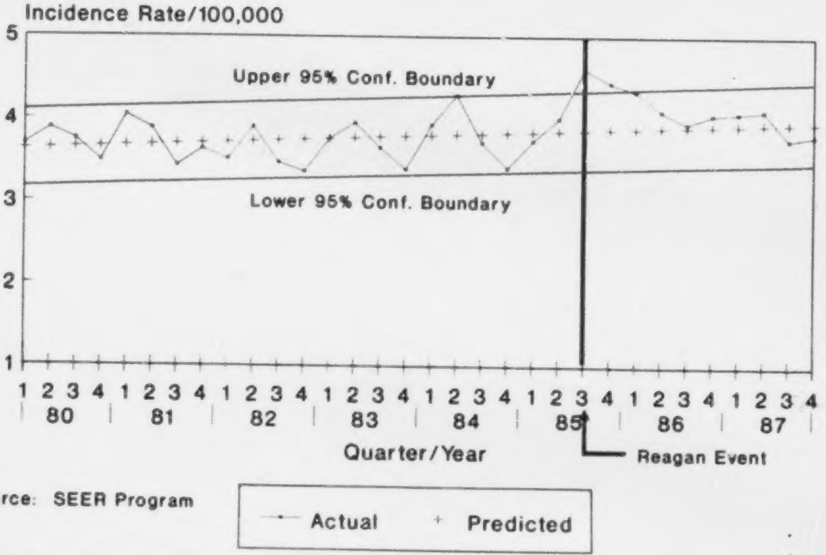
Figures 2 and 3 show age-adjusted quarterly incidence rates for colorectal cancer for the period 1980–1986, for early and advanced disease, respectively.[9] Each figure shows actual rates, rates predicted from an ordinary least squares regression (using data from 1973 to Second Quarter, 1985), and 95% prediction confidence intervals (also based on the data only from 1975 to Second Quarter, 1985). The 95% prediction confidence intervals have the following meaning for the period Third Quarter, 1985 through 1987: Given the information that was known

---

[7] The second largest increase, 22.6% for 1987, has to be interpreted with caution because it comes from a broader sample than the data for the period 1983–86. See sources listed at the bottom of Table 1.

[8] A transient increase in the measured incidence of early stage cancer is expected to follow a sudden increase in screening activity simply because a reservoir of "prevalent" cases of previously undetected early disease will exist in the population. This increase in "apparent incidence" is without any public health significance, however, unless it represents cases which, in the absence of early detection, would develop into more advanced, less treatable disease. Thus, a necessary, although not sufficient, condition for a public health impact to exist is that the initial increase in the measured incidence of early stage disease should be followed by a decrease in the measured incidence of advanced stage disease. Given the inherently noisy nature of cancer incidence data, this type of effect is not always easy to detect, even when it is actually present.

[9] The incidence of "early" colorectal cancer was defined as the incidence of "local stage" disease as defined and reported by the SEER cancer registry system. The incidence of "advanced" colorectal cancer was defined as the combined incidence of "regional stage" and "distant stage" disease as defined and reported by SEER. The computerized SEER records were accessed through NCI's Cancer Incidence and Surveillance System (KISS) software system.



Source: SEER Program

•One primary or first of multiple only

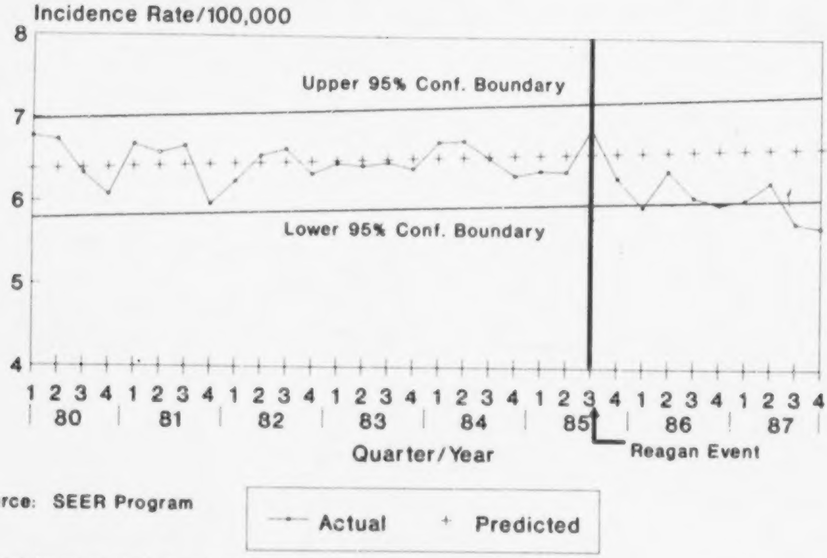
Figure 2. Colorectal Cancer, Early Disease

for the period 1973 to 1985 we would forecast, with a 95% level of confidence, that the rates for 1985 and 1987 would lie somewhere within the prediction confidence interval. If the actual rates for 1985 or 1987 lie outside of the prediction confidence interval we conclude that there is a good chance (19 out of 20) that this point does not fit the expected trend.[10]

It can be seen in Figure 2 that the incidence rate for early colorectal cancer lies well above the predicted trend for the period immediately following July, 1985. For the third and fourth quarters of 1985, the incidence levels lie above the upper bound of the 95% prediction confidence interval. For eight consecutive quarters following July, 1985, the incidence levels lie above the predicted trend, a result that is highly unlikely to occur by chance (a one in 256, or 0.4% chance). From Figure 3 it can be seen that there is a decline in the incidence of advanced disease, beginning in the fourth quarter of 1985. The incidence of advanced disease is on or below the lower bound of the 95% prediction confidence interval in five out of the eight quarters of 1986–87. Beginning with the fourth quarter of 1985, it is below the predicted trend for nine consecutive quarters. These observed relationships are consistent with an early detection effect. Some of the "excess" early

---

[10] The ordinary least square regressions and the prediction confidence intervals shown in Figures 2 and 3 were calculated using statistical software from SAS Industries, Inc.



Source: SEER Program

•One primary or first of multiple only;

Figure 3. Colorectal Cancer, Advanced Disease

disease detected in 1985 might have otherwise developed into advanced disease that would have been detected upon presentation of symptoms in 1986–87. The incidence data suggests that since this disease was "caught" more frequently in the early stage in 1985 there was less advanced stage disease to be found in 1986–87.

In summary, the temporal pattern observed for the incidence of early and advanced colorectal cancer following President Reagan's colon cancer episode is suggestive, albeit not conclusive, evidence that the subsequent increased public interest and early detection activity resulted in a genuine public health benefit.

ESTIMATES OF BENEFITS AND COSTS OF THE PRESIDENTIAL EFFECT

The magnitude of this possible benefit can be estimated from data on survival rates by stage-at-diagnosis of colorectal cancer. Since survival prospects are more favorable for early compared to advanced disease, the relative shift from advanced to early disease post-July, 1985, might be reflected in increased average survival and, potentially, reduced mortality from colorectal disease. Unfortunately, given our current state of knowledge about the natural history of colorectal disease and the characteristics of early detection techniques, we can estimate these possible benefits only with a wide margin of uncertainty.

Table 2 shows how the shift from advanced to early disease between

Table 2. Estimate of Survival Benefits Due to Presidential Effect

Years After Diagnosis	Cases Diagnosed in 1984	Cases Diagnosed in 1986		
		Upper Bound[a]	Middle[b]	Lower Bound[c]
	Number of People Still Alive (Out of 140,000)			
1	111,241	112,536	111,888	11,241
2	97,086	98,996	98,041	97,086
3	87,992	90,170	89,145	87,992
4	81,902	84,231	83,066	81,902
5	77,928	80,336	79,171	77,928
		Life Years Saved (1986 Cases Compared to 1984 Cases)		
1		1,295	647	0
2		1,910	955	0
3		2,178	1,153	0
4		2,328	1,164	0
5		2,408	1,243	0
Total		10,120	5,163	0

[a] Assuming no screening bias; e.g., relationship between observed stage and survival is identical for patients diagnosed in 1984 and 1986.

[b] Assuming 50% screening bias; e.g., only 50% of the increase in observed early stage disease actually results in true survival gains.

[c] Assuming 100% screening bias; e.g., none of the increase in observed early stage disease actually results in true survival gains.

1984 and 1986 *could* translate into this kind of survival benefit.[11] In 1984 the relative proportions of early and advanced disease were 36.8% and 63.2%; in 1986 these respective proportions were 40.2% and 59.8%. Based on these proportions and survival data for early and advanced disease, Table 2 shows the average survival prospects for the estimated 140,000 colorectal cancer cases diagnosed in 1986 and the hypothetical prospects of this same group if it had been diagnosed in 1984 (very similar results would be obtained by comparing 1987 to 1984). Three estimates are given in Table 2: the "upper bound" estimate assumes that 100% of the increase in observed early stage disease actually results in improved survival; the "lower bound" estimate assumes that none of the increase in observed early stage disease actually results in improved survival; and the "middle bound" estimate assumes that 50% of the increase in observed early stage disease actually results in improved survival. The resulting range, along the interval from zero to 10,120 person-years, is an estimate of the benefit which might accrue as a result of the increased early detection activity elicited by the publicity surrounding the President's colon cancer episode. It cannot be stressed too strongly that the numbers in Table 2 are the product of a hypothetical calculation, not observed data, using a range of assumptions about uncertain factors. This is a common and almost unavoidable approach in studies of cancer screening.[12] The most definitive evidence of the efficacy of specific cancer screening techniques would be the results of prospective, randomized, controlled clinical trials. Unfortunately, such evidence for colorectal screening will not be available for several years, at the earliest. In the absence of such data, follow-up studies of the Presidential Effect using actual survival and mortality data over the next few years could provide suggestive evidence of a general effect.

---

[11] The estimates in Table 2 are based on a simple model in that the effect of initially picking up prevalent, as opposed to incident, cases by screening is ignored.

Since public health benefits rarely come without a cost, it would be disingenuous not to report an estimate of the cost along with the benefit. On the cost side, while the focused media campaign can be taken as a windfall, the economic cost of elevated levels of proctoscopic examinations, and very likely fecal occult blood tests, is real. The cost to Medicare of the 28% increase in proctoscopic procedures from 1984–1985 amounted to about $90 million. Projected to the entire population age 50 and over[13] this cost increase would come to $196 million. Adding the increase in FOBT sales the total increase is approximately $200 million.

The purpose of this section has been to provide an order-of-magnitude estimate of the possible consequences of the Presidential Effect apart from mere attitudinal change. While our estimates are approximate and uncertain it is clear that these consequences are nontrivial.

---

[12] This uncertainty is a well-known issue in the literature on cancer screening. For example, if detecting cancer in a patient earlier than otherwise does not actually effect the true course of disease in that patient, then "lead-time" bias will result in the kind of calculation performed for the "upper bound" estimate in Table 2 being an overly optimistic estimate of screening benefits. Unfortunately, there is not a scientific consensus on how important lead-time, and other sources of bias, are in the case of colorectal cancer. Thus, the upper, middle, and lower bound estimates presented in Table 2 must be considered equally probable until more definitive knowledge is available (Morrison, 1985; Prorok, 1989).

[13] Using population data from the Bureau of the Census, 1988.

Conclusion

The evidence presented in this paper is only suggestive, but the temporal relationships between President Reagan's episode of colon cancer and the attendant publicity, the increase in phone calls to CIS related to colorectal cancer, the increased utilization of proctoscopy and FOBT, and the shift in the incidence rates of early and advanced colorectal cancer from predicted trends are all consistent with a potentially beneficial change in public health behavior induced by a high intensity media dissemination campaign.

It is important to appreciate the time lag effects involved in cancer statistics. First, there is a delay between the advent in the increase of early detection behavior and the consequent decrease in the incidence of advanced disease. Second, there is a delay in the change in incidence and the collection and reporting of cancer incidence data. Thus, we had to wait over three years from the date of President Reagan's colon cancer before the data reported in this paper were available. It is still too early to determine whether the favorable shift from advanced to early stage disease will persist past 1986. Determination of this will be important because it will help us gauge if the persistent but modest long-term increase in interest in colorectal cancer reflected in the frequency of CIS calls has a significant public health consequence or if only the dramatic but short-lived "spike" in interest immediately following the event was important. We will also have to wait a few more years before reliable mortality data are available. Only then, will it be possible to determine whether the reported favorable stage shift will actually be translated into a real benefit to the public's health.

References

DeVita, Vincent T. (1986)
"Interleukin-2, responsible reporting of an emotional issue." SIPIscope (Summer).
Doherty, Valerie R., and Rona M. Mackie (1988)
"Experience of a public education programme on early detection of cutaneous malignant melanoma." British Medical Journal 297:388–391.
Fink, Raymond, Ruth Roesner, Wanda Venet, Philip Strax, Louis Venet, and Mortimer Lacher (1978)
"Effects of news events on response to a breast cancer screening program." Public Health Reports 93:318–327.
Flay, Brian R. (1987)
"Mass media and smoking cessation: A critical review." American Journal of Public Health 77:153–160.
Heath, C. W., and D. J. Fink (1989)
"Trends in screening mammograms for women 50 years of age and older—behavioral risk factor surveillance system, 1987." Morbidity and Mortality Weekly Report 38:137–140.

Kinosian, Bruce P., and John M. Eisenberg (1988)
"Cutting into cholesterol, cost-effectiveness alternatives for treating cholesterolemia." Journal of the American Medical Association 259:2249–2254.

Levy, Alan S., and Raymond C. Stokes (1987)
"Effects of a health promotion advertising campaign on sales of ready-to-eat cereals." Public Health Reports 102:398–403.

Morrison, Alan S. (1985)
Screening in Chronic Disease. New York: Oxford University Press.

Office of Technology Assessment (1987)
Breast Cancer Screening for Medicare Beneficiaries: Effectiveness, Costs to Medicare and Medical Resources Required. Washington, DC: U.S. Government Printing Office.

Prorok, Philip (1989)
Personal communication, Philip Prorok, Chief, Screening Section, Biometry Program, Division of Cancer Prevention and Control, National Cancer Institute.

Rensberger, Boyce (1985)
"Reagan tumor removed; no cancer evident." Washington Post, 14 July.

Stoler, Peter (1985)
"What the diagnosis means: The facts and fears about colorectal cancer." Time 29 July, 22.

U.S. Bureau of the Census (1988)
Statistical Abstract of the United States. Washington, DC: U.S. Government Printing Office.

Warner, Kenneth E. (1987)
"Television and health education: Stay tuned." American Journal of Public Health 77:140–142.

A TALE OF TWO QUESTIONS: BENEFITS OF ASKING MORE THAN ONE QUESTION

ELIZABETH F. LOFTUS
MARK R. KLINGER
KYLE D. SMITH AND
JUDITH FIEDLER

Abstract In the last two months have you had a physical examination? The ability of 660 people to accurately report on their own recent health procedures was assessed by comparing the recollections of patients to information extracted from their medical records. A major finding was that when people were asked whether they had a specific procedure in a given reference period (e.g., physical exam within the last two months), overreporting occurred. A simple technique reduced overreporting: The technique entailed asking people about the procedure twice, first in connection with a different reference period (e.g., physical exam within the last six months) and then in connection with the reference period of interest (e.g., physical exam within the last two months). Our data support a "Precision" hypothesis: The advantage of the two–time frame questioning procedure is that it conveys to the respondent that the interviewer wants greater precision in dating than the single–time frame question implies.

"In the last two months, have you had a physical examination? How about an eye exam?" When people try to answer questions like these,

ELIZABETH F. LOFTUS is Professor of Psychology at the University of Washington. MARK KLINGER and KYLE SMITH recently received their Ph.D.s from that department. JUDITH FIEDLER is a Research Associate at Group Health Cooperative of Puget Sound. This research was supported by Grant Number HS 05521 from the National Center for Health Services Research and Health Care Technology Assessment. We thank Peggy Schnell for her invaluable assistance in chart abstracting, and Don Johnson for his many hours of assistance with data analysis. Liz Fries provided important support throughout many phases of this project. Without our tireless student research assistants, this project would have taken three times as long, and for this help we thank: LeAnn Davis, Jackie DuBois, Jeff Joyce, Marla Levin, Kris Morgan, Curt Small. All correspondence should be addressed to Elizabeth Loftus, Psychology Department, University of Washington, Seattle, Washington 98195.

Public Opinion Quarterly Volume 54:330–345 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

they are relying on retrospective memory and they often make mistakes (Baddeley, 1979; Strube, 1987). Sometimes they forget a procedure, and erroneously say "no" when they should say "yes." But forgetting is not the only error to worry about. Sometimes people recall an earlier episode (from, say, four months ago) as if it occurred more recently than it did. When a person reports an eye exam as having occurred during the previous two months when it actually occurred four months ago, a "forward telescoping" error has occurred. The data collected by survey researchers can be rendered misleading if forward telescoping is significant. (See also Garofalo and Hindelang, 1977; Loftus and Marburger, 1983.) Forward telescoping has been found to be a more serious problem than, say, backward telescoping, or the recalling of events as being further in the past than they actually were (Means et al., 1987; Sudman and Bradburn, 1973).

Bounded interviews are one solution to the problem of forward telescoping. In a bounded interview, the interviewer has the responses that were obtained from a prior interview conducted six months before the current interview. If the respondent in the current interview reports a recollection similar to one reported previously, the interviewer can probe further to determine whether the recollection now being reported is the same or is a new one, and if it is the same, it can be excluded from the current interview data. This technique has proven successful in surveys of household alterations and repairs (Neter and Waksberg, 1984). A major disadvantage of this technique is that it requires at least two interviews, and the data from the initial interview are used only to provide bounding information for subsequent interviews. Thus, it is wasteful of data.

The bounded interview technique was modified by Sudman, Finn, and Lannom (1984) for use in a single interview. These investigators first asked respondents about an earlier period and then about a more current period, all within a single interview session. This bounded recall procedure reduced telescoping in reports of health behavior and of snack food expenditures. While Sudman et al. were heartened by the apparent success of their simple technique, they explicitly warned that simultaneous validation information was essential to understanding the impact of bounding on errors of recall (p. 524).

A similar technique for potentially reducing forward telescoping was reported by Crespi and Swinehart (1982). Respondents were asked which of several medical procedures they had (blood pressure checked, physical exam, eye exam, etc.) in the past two months. Some respondents were first asked whether they had each procedure during the past six months, and then were asked the two-month question. When the six-month question was asked first, affirmative responses to the two-month question averaged 12% less than when the two-month question was asked alone. For example, when asked the six-month question first, 11% claimed to have had an eye exam in the last two months, 32% claimed to have had their blood pressure checked, 20% claimed to have had a physical exam. However, if the two-month question was asked alone, these figures were higher: 23%, 48%, and 32% respectively. In short, the two–time frame questioning procedure produced quite different estimates of the extent to which respondents engaged in the various health-related activities.

While the two–time frame procedure reduced the number of affirmative responses, did it lead to more accurate reporting? Crespi and Swinehart did not verify respondents' reports, but they had several reasons for believing that the lower reports were more accurate. Nonetheless, like Sudman et al. (1984), Crespi and Swinehart appealed to future researchers to reinvestigate their method with personal memories that could be verified.

It is natural to ask why people report fewer activities within a two-month period if they are first asked about activities within six months. There are a number of hypotheses. First, it is possible that events from three to four months ago are "captured" by the six-month question and thus do not need to be reported in response to the two-month question. For example, when asked the two-month question, Mary Smith remembers that she went to a wedding a few months ago and was struck in the eye by the bride's flying bouquet. She reports the eye exam that resulted from this mishap. If, however, she had first been asked the six-month question, she would have presumably already reported the exam. Now in response to the two-month question, she has further time to contemplate the event and realize it occurred more than two months before. Second, the need to demonstrate a socially desirable concern for health matters can be satisfied when respondents answer the six-month question and they do not need to demonstrate this concern again for the two-month question. Put another way, respondents want to be able to tell you about these sorts of events; they feel they are assisting the interview when they can provide a positive example of what you are asking about, and the six-month question gives them a chance to do this. Finally, it is possible that the two–time frame procedure conveys that the interviewer wants greater precision in dating than the single–time frame question might imply. When Mary Smith is asked the two-month question about eye exams after having just been asked the six-month question, she realizes that the interviewer cares about the precise dating of her experiences.

Our experiment was designed, in part, to explore the two–time frame procedure as a potentially useful means of increasing the accuracy of reported memories. In addition, we hoped to learn something about why people were reporting fewer activities when asked two questions rather than one. Crespi and Swinehart showed that people reported fewer activities in the two-month period when asked about a longer reference period first. Would people also report fewer activities when asked about a shorter reference period first? If the percentage reporting an activity were reduced when a shorter reference period came first, what explanations could account for this result? The precision hypothesis would predict that respondents who were first asked about one month, and then about two months, might sense a greater demand for precision and respond more accurately to the two-month question than respondents asked only about two months. The telescoping or social desirability explanations for the two–time frame effect would predict that asking about a shorter reference period first would not improve accuracy over a single–time frame question.

In sum, the purpose of this research was to determine whether a two–time frame procedure yields more accurate and complete reporting than a single–time frame procedure. Additionally, we aimed to distinguish between several general classes of explanations for reduced overreporting when two time frames are used.

Method

Our research was carried out in collaboration with Group Health Cooperative (GHC) of Puget Sound, the largest consumer-governed health maintenance organization in the United States. The Center for Health Studies at GHC conducts telephone surveys of Group Health clients on a regular basis, using standardized questionnaires from which interviewers read. By selectively manipulating the item content within one such survey, we conducted several experiments. Each subject participated in one or two of the individual experiments. In this article, we report only on the results from the two–time frame study.

SUBJECTS

Prospective survey participants were randomly selected by computer from GHC membership rolls. Letters were mailed to approximately 1,200 members asking for their participation in a consumer survey. Of these, 950 members completed the survey, yielding a completion rate of 79.2%. The sample was randomly drawn within stratified geographical regions. Of the members who participated in the survey, 94% gave permission for their medical records to be reviewed.

A good deal of demographic information was gathered from respondents, only a portion of which was relevant to this study. Of those who participated in the survey, 37% were males and 62% females (1% unknown). The participants who were interviewed ranged in age from 14 to 82, with a mean of 42.5 years. 15% were over 65. Adults and young people 14 to 17 years old responded for themselves. Data for children 2 to 13 years old were obtained by interviewing parents as proxies for the children, and 19.3% of the sample consisted of parental proxy reports.

In this two–time frame study, 739 members participated and granted permission to have their medical records checked.[1] Permission was sought at the end of the survey.

PROCEDURE

All participants were interviewed during March, 1986. Each respondent who agreed to participate was telephoned and asked a series of approximately 50 questions (some of which had subquestions). The questions were part of a routine consumer opinion survey conducted by GHC to gather opinions and evaluations of GHC services and to provide information to assist in the planning of health care programs and facilities. The survey lasted approximately 30 minutes. Specific questions pertinent to this experiment were embedded in the beginning of this questionnaire. In order to incorporate the manipulations of item content required by the experiment, we constructed three different versions of the survey questionnaire. Participants were randomly assigned to one version.

All versions of the survey began in the same way. The members were asked how many years they had been enrolled at GHC for health care, and the specific medical center to which they usually went. After a few more preliminary questions, they were asked the critical questions that constituted the two–time frame experiment. Respondents were asked whether they had any of a list of procedures (e.g., blood pressure reading) done under their GHC coverage. Several of the procedures on this list were routine, while others were ones that the GHC staff were particularly interested in. The three versions of the questionnaire differed in terms of the reference period about which respondents were asked. In the 2-6 condition, respondents were first asked about procedures utilized over the previous two months, and then about procedures over the previous six months. Because respondents could not anticipate the sequence of questions, responses to the two-month question appearing first in the 2-6 sequence may be considered equivalent to responses to a two-month question asked alone. Adult respondents who were over 18 years ($N = 180$) were asked about 7 different procedures:

---

[1] This research was part of a larger project on enhancing memory for health events. The details of the larger project can be found in Loftus, Smith, Klinger, and Fiedler, 1989.

> "During the past two months, since (date), have you had any of the following procedures done under your GHC coverage?
> Blood pressure reading?
> Test of blood in your stool?
> Mammogram—x-ray of breast (asked only of women over 40 years)?
> Pneumococcus vaccine—flu vaccine (asked only of people over 65 years)?
> Breast exam by a doctor or nurse (asked of women only)?
> Pap smear test for cancer (asked of women only)?
> Had a new prescription filled at a GHC pharmacy?"

Next, the respondent was asked, "During the past six months, since (date), have you had any of these procedures at GHC?" and the procedures were repeated to them.

Adolescents between 14 and 17 years old were asked about a different set of procedures. Also parents responding for their children below age 13 were asked about this set of "child" procedures ($N=62$).

> "During the past two months, since (date), have you (has your child) had any of the following procedures at GHC?
> Physical examination?
> Urine test?
> Vision test?
> Hearing test?
> Had a new prescription filled at a GHC pharmacy?"

Next, the respondent was asked, "During the past six months, since (date), have you (has s/he) had any of these procedures at GHC?" and the relevant procedures were repeated to them.

In the 6-2 condition, respondents were first asked about procedures utilized over the previous six months, and then about procedures over the previous two months. Adult respondents who were over 18 years old ($N=180$) were asked about the seven "adult" procedures listed earlier, using the six-month reference period. Next, if they answered yes, they were asked about the relevant procedures using the two-month reference period. Adolescents who were 14 to 17, and parents reporting for their young children ($N=73$), were asked about the five "child" procedures listed earlier, using first the six-month and then the two-month reference period.

In the 1-2 condition, respondents were first asked about procedures utilized over the previous month, and then about procedures over the previous two months. Adult respondents who were over 18 years old ($N=188$) were asked about the seven "adult" procedures listed earlier, using the one-month, then the two-month reference period. Adolescents aged 14–17, and parents reporting for their young children (*N* = 56), were asked about the five "child" procedures using first the one-month, and then the two-month reference period.

Results

For this analysis, we considered responses to the two-month question appearing first in the 2-6 condition to be equivalent to responses to a two-month probe alone. We then compared these responses to those received from the 6-2 and from the 1-2 respondents, in terms of their recollections of the two-month reference period.

Do people report more procedures when asked about a single time frame (e.g., two months) than when asked about that same frame, preceded by another time frame (e.g., six months, then two months)? One answer to this question can be found in Table 1, showing the percentage of respondents who reported a procedure within the last two months.

First, compare the 2-6 condition to the 6-2 condition. For 9 out of 12 separate procedures, there were fewer visits reported in the 6-2 condition. That is, for 9 out of the 12 procedures, asking the six-month question first reduced the reporting of visits that had occurred in the previous two months, as Crespi and Swinehart had shown.

Table 1. Percentage Reporting a Procedure in the Last Two Months

		Condition		
	N	6-2	2-6	1-2
Blood pressure reading	542	31.8	29.8	32.1
Colon cancer test	540	5.1	5.6	7.1
Mammogram (women over 40)	144	17.1	5.0	0.0
Pneumococcus vaccine (over 65)	62	5.0	0.0	0.0
Breast exam (women)	301	18.3	18.7	12.0
Pap smear (women)	301	8.3	15.6	12.0
Rx filled	543	24.5	30.4	33.9
Child physical exam	190	9.8	15.3	13.7
Child urine test	187	6.7	12.3	5.9
Child vision test	189	0.0	6.9	11.8
Child hearing test	189	0.0	7.0	7.8
Child Rx filled	188	22.6	28.6	28.6

Next compare the 2-6 condition to the 1-2 condition. There is no consistent pattern showing fewer reports in the 1-2 condition, that is, when a shorter time frame precedes the time frame of interest.

Another comparison can be made to shed light on whether a sequencing effect would appear when a shorter time frame precedes the time frame of interest. We can compare the responses to the six-month question in the 2-6 condition to the responses to the six-month question when asked first (as it is in the 6-2 condition). The percentage of respondents who reported a procedure within the last six months is shown in Table 2.

Compare the 6-2 to the 2-6 condition. For 11 out of 12 procedures, the 2-6 condition led to fewer reported visits during the six-month reference period. Thus, we have another instance in which having respondents report activities in two time frames reduces the proportion reporting the activity in the second time frame.

In sum, there appears to be a modest reduction in the percentage who report a visit over a two-month period, when these respondents are first asked about a six-month period. As for whether a shorter reference period preceding the critical one has a similar effect, the results are mixed. Answering about a two-month period first reduced reporting in a six-month reference period; however, answering about a one-month period first did not have a reliable effect on answers for a two-month reference period.

One problem with these data are that we have no guarantee that the

Table 2. Percentage Reporting a Procedure in the Last Six Months

		Condition	
	N	6-2	2-6
Blood pressure reading	352	61.9	57.3
Colon cancer test	353	21.2	17.6
Mammogram (women over 40)	102	23.8	9.5
Pneumococcus vaccine (over 65)	43	30.0	25.0
Breast exam (women)	198	36.9	35.2
Pap smear (women)	196	29.8	28.4
Rx filled	357	49.1	47.2
Child physical exam	134	43.5	32.8
Child urine test	132	23.3	17.5
Child vision test	133	15.9	12.3
Child hearing test	132	14.8	11.1
Child Rx filled	132	48.3	44.4

actual numbers of visits are the same in the three conditions. To determine more meaningfully whether one condition leads to more overreporting than another, we must examine reports that can be verified. For this reason, we extracted information from the medical records of our participants.

ANALYSIS OF MEDICAL RECORDS

The first step in the medical record check was to order the charts of the respondents who had granted permission to have their records checked. Abstracting of charts was done by research assistants who were trained by GHC. For each participant, we obtained from the medical records information on several specific procedures (e.g., blood pressure reading) that the participant had had during the previous six months. This information was generally found in lab reports, radiology reports, or other parts of the medical chart. The abstracters recorded whether the respondent did or did not have each procedure. For each procedure found, the date was also noted. Once we had the survey and the medical record data, they were matched to determine the accuracy of recall.

A word of caution is in order here. We will talk about errors of recollection, but we must acknowledge that some of these "errors" could be due to problems in the medical records (Martin, 1987). Analyses of medical records have shown the occasional inaccurate or missing piece of information. While GHC has an exceptionally advanced method of record keeping, it still cannot be assumed that all discrepancies between the recollections and the medical records are errors of memory. For the analyses reported here, we will assume that the records are accurate, even if not perfect.

Accuracy of Recall. Of the 739 persons who participated in this experiment, we were able to verify medical information of 660 people. The analyses described in this section pertain only to the 660 verified cases.

For 10 of the 12 procedures, we computed three scores. From the medical records, we found the percentage who actually had the procedure during the previous two months. We compared this to the percentage who reported having the procedure during the previous two months. We then calculated the extent of overreporting by subtracting actual procedures from recalled procedures. These computations are shown in Table 3, for respondents who recalled over the two-month reference period. The *N*s refer to the number of people who were asked about the procedure for whom verification could be obtained.

The data in Table 3 reveal consistent overreporting. It ranges from 2% for the flu vaccine to over 6% for the breast exam and blood

Table 3. Overreporting for Two Months

	N	Recalled	Actual	Overreported	Ratio
Blood pressure reading	486	31.3	25.1	6.2	1.25
Colon cancer test	487	6.0	2.1	3.9	2.86
Mammogram (women over 40)	119	7.6	3.4	4.2	2.24
Pneumococcus vaccine (over 65)	49	2.0	0.0	2.0	—
Breast exam (women)	265	16.2	9.8	6.4	1.65
Pap smear (women)	266	12.0	7.1	4.9	1.69
Child physical exam	171	13.0	10.1	2.9	1.29
Child urine test	168	8.2	4.7	3.5	1.75
Child vision test	171	5.8	2.3	3.5	2.52
Child hearing test	170	4.7	1.2	3.5	3.92

pressure check. The median amount of overreporting across procedures was 3.7%. The extent to which overreporting is a serious problem can be more readily determined by examining the ratios of reported procedures to actual procedures, listed in the last column of Table 3. Note that it was common for reported procedures to be twice as great as actual procedures, and in at least one instance, the reported procedure was almost four times as great as actual. In Table 4, the same data are presented for the six-month reference period. Here even

Table 4. Overreporting for Six Months

	N	Recalled	Actual	Overreported	Ratio
Blood pressure reading	312	59.6	44.6	15.0	1.34
Colon cancer test	315	19.4	7.0	12.4	2.77
Mammogram (women over 40)	84	16.7	9.5	7.2	1.76
Pneumococcus vaccine (over 65)	36	27.8	8.3	19.5	3.35
Breast exam (women)	172	36.0	17.4	18.6	2.07
Pap smear (women)	173	28.9	16.2	12.7	1.78
Child physical exam	117	38.5	26.5	12.0	1.45
Child urine test	119	21.0	10.9	10.1	1.93
Child vision test	121	14.0	6.6	7.4	2.12
Child hearing test	118	13.6	5.9	7.7	2.31

more massive overreporting is apparent. Overreporting ranges from a low of about 7% for the mammogram to a high of nearly 20% for the flu vaccine. The median amount of overreporting across procedures was 12.5%.[2] Again, it was common for reports of a given procedure to be twice as great as the actual number of procedures.

For all procedures, we found massive overreporting. We can now ask whether the overreporting problem is different as a function of being asked about a single time frame versus two time frames. To see how our analyses were done, consider the blood pressure check for which we provide complete data for the two-month reference period in Table 5. In the 2-6 condition, 29.8% recalled having a blood pressure check, which was quite close to the percentages recalling this procedure in the other two conditions (31.8% in the 6-2 condition and 32.1% in the 1-2 condition). However, when one subtracts the actual percentages from the recalled percentages in each of the three conditions, it becomes evident that overreporting is a more serious problem in the 2-6 condition than in the other two conditions. In the 2-6 condition, the overreporting of blood pressure checks is 10.5% while in the 6-2 condition it is only 3.1% and in the 1-2 condition it is 4.7%. Thus, there is overreporting in all three conditions, but it is greatest when the two-month reference period is asked about first.

Another observation from Table 5 concerns the false alarms, that is, cases in which people recalled having their blood pressure checked, but their record did not confirm this. The false alarm rate was 50% in the 2-6 condition, but was lower when another question preceded the one about the two-month reference period (32% in the 6-2 condition and 25.6% in the 1-2 condition). In short, false alarms occurred much less often with the two–time frame approach.

In contrast to false alarms, the pattern for misses was different. Misses occurred when people had had their blood pressure checked in the previous two months, but claimed they had not. In the 2-6 condition, 6.2% of actual blood pressure checks were not recalled by respondents, as compared to a somewhat higher number in the two–time frame conditions (10.3% in the 6-2 condition and 7.0% in the 1-2 condi-

---

[2] Notice that the prescription data are not included in the above analyses. They were deliberately omitted because of the following problem. The respondents were asked whether they had filled a *new* prescription (for themselves or their child) during the reference period. In retrospect, this wording was unfortunate. Suppose a respondent got a prescription that could be refilled once. If the doctor reissued the prescription, should it be classified as new or not? A second problem arose with the medical record data. The medical record data provided dates on which prescriptions were filled, but did not distinguish between old and new prescriptions. This meant that we had numerous instances in which respondents told us that they had not had a new prescription filled, but the records told us that prescriptions had been filled, leading to massive but misleading underreporting. Given these problems, the only sensible thing to do was ignore the prescription data.

Table 5. Blood Pressure Check in Previous Two Months

	Medical Records		
Participants' Response	Had Test	Did Not	Total
	Two-Month Question First		
Yes, I had my B.P. checked.	24	24	48
	50.0%	50.0%	29.8%
No, I did not.	7	106	113
	6.2%	93.8%	70.2%
Total	31	130	161
	19.3%	80.7%	100.0%
	Two-Month Question Following Six-Month Question		
Yes, I had my B.P. checked.	34	16	50
	68.0%	32.0%	31.8%
No, I did not.	11	96	107
	10.3%	89.7%	68.2%
Total	45	112	157
	28.7%	71.3%	100.0%
	Two-Month Question Following One-Month Question		
Yes, I had my B.P. checked.	38	16	54
	70.4%	29.6%	32.1%
No, I did not.	8	106	114
	7.0%	93.0%	67.9%
Total	46	122	168
	27.4%	72.6%	100.0%

tion). Thus there were slightly more misses with the two–time frame approach.

We presented the blood pressure data as our illustrative example not only to show how our analysis was done, but also because by a large margin more respondents either had this procedure or recalled the procedure than was the case with any other procedure tested in this study. The most important finding from the blood pressure data was that overreporting was decreased (and thus reporting more accurate) when two time frames were used compared to a single time frame. But the beneficial effect of the two–time frame procedure is not limited to the blood pressure data.

A similar pattern can be seen in the breast exam data. In the two-month condition, overreporting was 6.4%. All three conditions showed overreporting for a breast exam. However, the overreporting was greater when the two-month reference period was asked about first (11.0% overreporting), compared to when it was asked about second (7.3% overreporting for the 6-2 condition, and 1.1% overreporting for the 1-2 condition). The false alarm rate was high in all conditions. Of those who claimed to have had a breast exam, 77% were false alarms in the 2-6 condition, whereas this figure was lower in the 6-2 condition (53%) and in the 1-2 condition (36%). Misses were lower than false alarms. Of those who actually had a breast exam, 43% failed to report it in the 2-6 condition, whereas the miss rate was lower in the two–time frame conditions (22% in the 6-2 condition and 30% in the 1-2 condition). In sum, overreporting occurred in all conditions, but was highest when the two-month reference period was asked about first. This condition also led to a higher false alarm and miss rate than the other two conditions in which the two-month reference period was preceded by an earlier question.

When all procedures, including the blood pressure and breast exams, are taken into account, the beneficial effects are revealed. To see this, examine Table 6, which presents summary rather than detailed data for all procedures.

Ten different procedures are included in Table 6. When we compare the two-month alone to the 6-2 condition, we find that for 7 of the 10 separate procedures, there was less overreporting in the 6-2 condition. Four out of 10 procedures were at least marginally better off in the 6-2 condition, and none were worse off. When we compare the two-month alone to the 1-2 condition, we find that for 7 of the 10 separate procedures, there was less overreporting in the 1-2 condition. Two out of 10 procedures were at least marginally better off in the 1-2 condition, and one was worse off. Finally, when we compare the six month alone to the 2-6 condition, we find that for 8 of the 10 separate procedures, there was less overreporting in the 2-6 condition. Four out of 10 separate procedures were at least marginally better off in the 2-6 condition, and only 1 was worse off.

In sum, if we now examine our data across procedures, we have available 30 separate comparisons of the single– versus the dual–time frame approach. For 22 of the 30 separate procedures, there was less overreporting in the dual–time frame conditions. Ten of 30 separate procedures were significantly or marginally better off in the dual–time frame than in the single–time frame condition, and two procedures were marginally worse off.

Table 6. Percentage Overreporting by Recall Condition

	Two-Months			Six-Months	
	2-First	6-2	1-2	6-First	2-6
Blood pressure reading	10.5	3.1***	4.7*	11.6	18.4#
Colon cancer test	2.5	4.5	4.7	16.1	8.0**
Mammogram (women over 40)	2.5	9.8	0.0	9.5	4.7
Pneumococcus vaccine (over 65)	0.0	5.0	0.0	30.0	6.2
Breast exam (women)	11.0	7.3	1.1***	19.0	18.2
Pap smear (women)	7.8	1.2**	4.4	9.6	15.9
Child physical exam	6.8	0.0**	3.9	17.7	8.7
Child urine test	3.5	3.4	3.9	15.0	3.5**
Child vision test	1.7	0.0	7.9#	8.0	5.3
Child hearing test	3.5	0.0**	7.8	9.9	5.5

NOTE: All statistical tests compare two–time frame conditions with single–time frame conditions.

$^{***}p < .01$. $^{**}p < .05$. $^{*}p < .10$. $^{\#}p < .10$, in opposite predicted direction.

Discussion

To reiterate our major result, when people were asked whether they had a specific procedure (e.g., flu shot) in the recent past, they overreported. The median amount of overreporting for the procedures we asked about in the six month reference period was approximately 13%. For many procedures, two to three times as many people reported having had a procedure done as actually did. Overreporting was more extensive over the six-month reference period than the two-month reference period. Yet even after only two months, the extent of overreporting was 4% (across all procedures).

Can we do anything to reduce the overreporting that occurs in, say, the two-month reference period? Crespi and Swinehart tried to reduce error in the reporting of procedures by asking about procedures in either a single–time frame or two–time frames. They found that if respondents first answered about, say, a six-month reference period and then a two-month reference period, the reporting was lower than if they asked about the two-month question first. They never verified reporting to determine whether the lower reporting meant more accurate reporting. We did. Like Crespi and Swinehart, we found less overreporting in the two-month reference period when respondents were first asked about six months.

There are several possible reasons for the beneficial effect of the 6-2 condition. The two–time frame procedure may work for social desirability reasons. If a respondent has the opportunity to mention a health maintenance procedure as having occurred during the longer reference period, the need to give a socially desirable response has been satisfied, and he or she does not need to mention it again during the shorter reference period. Another explanation is that the two–time frame procedure conveys to the respondent that the interviewer desires greater precision in responding. The precision hypothesis derives a modicum of support from the observation that the 1-2 condition also produced more accurate responding. Moreover, when asking about a six-month reference period, the 2-6 condition produced less overreporting than asking about six months straight away. Thus, we have two demonstrations that having a shorter reference period prior to the period of interest increases the accuracy of reporting. The benefits of asking about two time frames appear to about the same degree no matter whether the first time frame is longer or shorter than the second one. The social desirability hypothesis has no particular explanation for the comparable benefits by asking first about a shorter time period.

It is not a new idea that respondents, when left to their own devices, often make careless and imprecise estimates of the occurrence of particular events (Cannell et al., 1977). It was suggested in one study, for example, as a possible reason for erroneous reporting of drug use (Bachman and O'Malley, 1981). We have speculated here that the two–time frame procedure may communicate to the respondent that a greater sense of precision is desired. It would be a relatively straightforward task for future research to test this more directly. One could ask the two–time frame respondents a question like "On a scale from 1 to 7, how precise an answer do you think we wanted? 1 = just a rough estimate, 7 = a very precise answer?" Our prediction is that the two–time frame respondents will choose a higher scale value than their single–time frame counterparts. If the precision hypothesis is correct, then other, perhaps even more efficient, methods of enhancing the perceived precision requirement could be constructed.

One difference between our procedure, and that of Crespi and Swinehart, is that we gave respondents a date that marked the beginning of the reference period. The date was given once, at the beginning of the list of procedures inquired about, and thus may have been ignored or forgotten by some respondents. Those who did not keep the date in mind while working through the list of procedures, may have not sensed that the question was demanding precision. The two–time frame procedure might have worked only for this subgroup. This train of thought leads to the speculation that the beneficial effect of the two-time frame procedure would have been even greater if the date had not been included.

References

Bachman, J. G., and P. M. O'Malley (1981)
"When four months equal a year: Inconsistencies in student reports of drug use." Public Opinion Quarterly 45:536–548.
Baddeley, A. (1979)
"The limitations of human memory: Implications for the design of retrospective surveys." Pp. 13–30 in L. Moss and H. Goldstein (eds.), The Recall Method in Social Surveys. London: University of London Institute of Education.
Cannell, C. F., K. H. Marquis, and A. Laurent (1977)
"A summary of studies of interviewing methodology." Pp. 1–78 in Vital and Health Statistics, U.S. Public Health Service, Series 2.
Crespi, I., and J. W. Swinehart (1982)
"Some effects of sequenced questions using different time intervals on behavioral self-reports." Paper presented at the American Association for Public Opinion Research, Washington, DC.
Garofalo, J., and M. J. Hindelang (1977)
An Introduction to the National Crime Survey. Washington, DC: U.S. Dept. of Justice.
Loftus, E. F., and W. Marburger (1982)
"Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events." Memory and Cognition 11:114–120.
Loftus, E. F., K. D. Smith, M. R. Klinger, and J. Fiedler (1989)
"Memory and mismemory for health events." Unpublished manuscript, University of Washington.
Martin, C. J. (1987)
"Monitoring maternity services by postal questionnaire: Congruity between mother's reports and their obstetric records." Statistics in Medicine 6:613–627.
Means, B., A. Nigam, M. Zarrow, E. F. Loftus, and M. S. Donaldson (1987)
"Autobiographical memory for health-related events: Enhancing memory for recurring incidents." Final Report. Alexandria, VA: HUMRRO.
Neter, J., and J. Waksberg (1964)
"A study of response errors in expenditures data from household interviews." Journal of the American Statistical Association 59:18–55.
Strube, G. (1987)
"Answering survey questions: The role of memory." In H-J. Hippler, N. Schwarz, and S. Sudman (eds.), Social Information Processing and Survey Methodology. New York: Springer-Verlag.
Sudman, S., and N. M. Bradburn (1973)
"Effects of time and memory factors on response in surveys." Journal of the American Statistical Association 68:805–815.
Sudman, S., A. Finn, and L. Lannom (1984)
"The use of bounded recall procedures in single interviews." Public Opinion Quarterly 48:520–524.

THE EFFECT OF MONETARY INCENTIVES AND FOLLOW-UP MAILINGS ON THE RESPONSE RATE AND RESPONSE QUALITY IN MAIL SURVEYS

JEANNINE M. JAMES AND
RICHARD BOLSTEIN

Abstract The joint and comparative effects of the use of monetary incentives and follow-up mailings were examined in a mail survey of suburban Washington, DC cable television subscribers. Four experimental groups received monetary incentives enclosed with the first mailing only ($0.25, $0.50, $1.00, or $2.00) and three follow-up mailings. These groups were compared with each other and against a control group that did not receive an incentive. The results indicated that the response rate from the first mailing increased significantly as the incentive amount increased from zero to $0.25, and from $0.25 to $1.00. Four mailings without an incentive produced a higher response rate than a single mailing with an incentive, but a *combination* of follow-up mailings and a $1.00 or $2.00 incentive produced a significantly higher response rate than an equivalent number of mailings without an incentive.

There was some evidence of intertreatment response bias. Larger monetary incentives tended to produce: (1) a greater degree of effort expended in completing the questionnaires, as measured by the number of short answers and comments provided, and the number of words written, and (2) comments that were more favorable toward the survey sponsor.

JEANNINE JAMES is President of American Research Company, Inc., Great Falls, Virginia. RICHARD BOLSTEIN is Associate Professor, Center for Computational Statistics, Department of Operations Research and Applied Statistics, George Mason University, Fairfax, Virginia. This study was conducted as a doctoral dissertation in Industrial-Organizational Psychology at George Mason University, Fairfax, Virginia, by Jeannine James, who wishes to thank Louis Buffardi, Director of the Industrial-Organizational Psychology Doctoral Program, and Theodore Gessner, Associate Professor of Psychology for their help respectively as chairman, and advisor on the dissertation committee. The authors also wish to thank Don Mathison, Vice President, Marketing and Programming, Media General Cable Company for providing the survey vehicle, and the Dean of the Graduate School, George Mason University, for providing a grant for the monetary incentives.

Public Opinion Quarterly Volume 54:346–361 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

Introduction

According to the mail survey literature reviews, the two most effective methods of increasing response rates are: (1) sending follow-up letters with replacement questionnaires, and (2) enclosing incentives with the initial mailing (Blumberg, Fuller, and Hare, 1974; Fox, Crask, and Kim, 1988; Heberlein and Baumgartner, 1978; Kanuk and Berenson, 1975; Linsky, 1975; Robinson and Agisim, 1951; Yu and Cooper, 1983). Moreover, monetary incentives have been shown to be significantly more effective than premiums (such as key chains or ballpoint pens) (e.g., Goodstadt, Chung, Kronitz, and Cook, 1977; Hansen, 1980).

But how do these two methods compare with one another, and are they more effective when used in combination than when used alone? During the past 50 years a vast quantity of controlled experimentation has produced some well-defined response maximization techniques (e.g., multiple follow-up mailings, personalization, enclosure of postage paid envelope, design of questionnaire that is easy to fill out) which consistently generate response rates of 70% or more (Dillman, 1978). In light of the effectiveness of these techniques, perhaps it is no longer necessary or useful to send monetary incentives.

A few studies have compared the response rates obtained with follow-up mailings versus prepaid monetary incentives, but with conflicting results. For example, Kephart and Bressler (1958) tested the effect of multiple incentive levels (1, 5, 10, and 25 cents) in a two-wave survey. They found that, although after the first mailing the 25 cent incentive was the only amount that generated a significantly greater response rate than the no-incentive condition, after the second mailing, the cumulative response rates did not differ significantly between the 25 cent and the no-incentive condition. On the other hand, a study by Goodstadt, Chung, Kronitz, and Cook (1977) found that a 25 cent incentive produced a significantly higher response rate than the use of no incentive even after three follow-up mailings to each group.

In a more recent study, Nederhof (1983) compared the use of a nonmonetary incentive, a ballpoint pen, with follow-up mailings and found that after one mailing, the group that received the ballpoint pen showed a significantly higher response rate than the no-incentive group. However, after one follow-up mailing to both groups, the response rate difference diminished to a nonsignificant amount, and after a second follow-up mailing, the difference was reduced even further. Nederhof concluded that "when methods that produce high response rates are used, the effect of nonmonetary incentives on response rates disappears after the first mailing" (Nederhof, 1983). Since monetary incentives are more potent, however, the question remains as to whether a monetary incentive response rate advantage would be diminished as quickly as the nonmonetary incentive.

According to the literature reviews, a monetary incentive as small as 25 cents is sufficient to produce a response rate significantly higher than a no-incentive control condition (e.g., Armstrong, 1975; Blumberg, Fuller, and Hare, 1974; Kanuk and Berenson, 1975; Linsky, 1975; Robinson and Agisim, 1951), and larger monetary incentives do not have a significantly greater effect (Armstrong, 1975; Kanuk and Berenson, 1975). However, the reviews on this topic are generally based on studies conducted between the 1940s and 1960s, and most of these studies have only tested the effectiveness of a 25 cent incentive versus a no-incentive control in one-wave mailings; only a few of the studies tested larger incentive amounts. So it remains unclear whether or not an incentive amount larger than 25 cents might be significantly more effective, and whether an amount larger than 25 cents might be necessary today to have a significant effect.

Another issue of importance is the question of whether incentives create a bias by attracting respondents who are demographically different from nonremunerated respondents, or by affecting the quality or quantity of the responses given. Most previous studies have found that the incentives have little effect on the sociodemographic composition of respondents (e.g., Furse and Stewart, 1982; Hansen, 1980; Robertson and Bellenger, 1978; Wotruba, 1966), but at least one study has found that incentives produce a disproportionately larger percentage of respondents from lower educational and occupational strata (e.g., Nederhof, 1983). Most studies have found that the answers given by respondents do not differ as a function of incentives (Goodstadt, Chung, Kronitz, and Cook, 1977; Hancock, 1940; Hansen, 1980; Heads and Thrift, 1966; Maloney, 1954; Mizes, Fleece, and Roos, 1984; Nederhof, 1983). However, response quality, defined as "the degree of effort and thought devoted to the questionnaire by the respondent who completes it" (Houston and Ford, 1976), has been found to be greater as a result of financial incentives (e.g., Godwin, 1979; McDaniel and Rao, 1980).

Another issue of importance is the question of the cost-effectiveness of using monetary incentives versus follow-up mailings.

OBJECTIVES

This study compares the mail survey response rates that result from using various monetary incentives enclosed with the first mailing, versus, and in combination with, the use of three follow-up mailings without incentives. The possibility of intertreatment response bias is examined by comparing differences in: (1) demographic characteristics of respondents, (2) answers and opinions given, and (3) quality and quantity of information provided. A cost-benefit analysis of each experimental condition is also performed. We begin by describing the research method, then present the results and discussion of (*a*) response rate, (*b*) cost analysis, and (*c*) response quality.

Method

A four-page mail survey questionnaire containing fixed-response and open-ended questions was sent to 850 cable television subscribers in Fairfax County, Virginia (a suburb of Washington, DC) in the fall of 1986. The objective was to obtain a profile of the cable television subscribers, as well as a profile of users and nonusers of a particular cable television service to be used for marketing and programming purposes by the cable television company sponsor.

The sampling frame lists households in order by subscription date. Since cable television service becomes available on a neighborhood basis, and since most neighborhoods in Fairfax are rather homogeneous with regard to demographics such as age, education, and income, a systematic random selection was used. The sample was then systematically assigned to a control group and four experimental conditions: each group consisted of one of the first five subscribers from the sample and every fifth thereafter. Subjects in the experimental groups received a monetary incentive of either $0.25, $0.50, $1.00, or $2.00 enclosed with the first mailing, while the control group did not receive a monetary incentive. One-half of the subjects in the $0.50 group received a fifty-cent piece, while the other half received two quarters. In the $2.00 group, one-half of the subjects received a $2.00 bill; the other half received two $1.00 bills (again, systematic random assignment was used).

The "Total Design Method" (Dillman, 1978) was used as a model for the survey procedures and the wording of the cover letters. The cover letters were the same for all groups, with the exception that the letter sent to the experimental subjects contained a P.S. that explained the incentive: "Although the enclosed $0.25 ($0.50/$1.00/$2.00) cannot possibly compensate you for your time and effort on our behalf, please accept this as a small token of our appreciation for your cooperation in completing and returning the enclosed questionnaire." Three follow-up mailings without monetary incentives were sent to the nonrespondents in all the groups at three-week intervals. For the third follow-up mailing, the nonrespondents in each of the five groups were randomly divided into two groups, a first class mailing was sent to one group and a certified mailing to the other.

Response Rate

RESULTS

An overall response rate of 91% (767/844) was obtained after the four mailings. (The sample size was reduced by six ineligibles.) After the first mailing the response rates in the fifty-cent piece group (62%) and the two quarters group (64%) were virtually the same. Similarly, the two $1.00 bills did not generate a significantly higher response rate (82%) than the one $2.00 bill (73% [$\chi^2(1, N = 170) = 2.17, p = .16$]). The certified mail procedure used in the final mailing did not generate a significantly higher response rate (28%) than first-class mail (19% [$\chi^2(1, N = 106) = 1.07, p = .30$]). However, it should be noted that the sample sizes were small for the comparisons made within the $0.50 and $2.00 incentive groups ($N = 85$), and in the certified and first-class mail procedures ($N = 54$ and $N = 52$), and the effect would have had to be very large to be detected by the tests of significance. These issues need to be investigated further in another study.

Due to the nonsignificant results between subgroups, the data from the fifty-cent group, the two-dollar group, and the fourth-mailing group are aggregated in the remainder of this paper.

FIRST MAILING

Following the first mailing there were highly significant differences in response rates between the incentive groups ($\chi^2[4, N = 844] = 25.69$, $p < .001$, see Table 1). There is a clear tendency for the response rate to increase with the incentive amount. The likelihood ratio test (Robertson et al., 1988:166–167) confirms a monotonic increasing response rate ($-2\ln\lambda = 26.07, p < .001$).

To discover the incentive amounts that have the greatest effect on response rate, we partitioned chi-squared (Agresti, 1984:12–15). Because of the monotonic relationship, all pairwise tests are one-tailed. Since the response rates for the $0.25 and $0.50 groups are indistinguishable, we combined them and tested the aggregate against the control group. We obtained a significant difference ($\chi^2[1, N = 505] = 3.57, p < .05$), indicating that the $0.25–$0.50 group has a higher response rate than the control group. Similarly, comparison of the $0.25–$0.50 group with the $1.00 group produced a significant difference ($\chi^2[1, N = 506] = 4.90, p < .02$). Finally, a comparison of the $1.00 versus $2.00 incentive produced nonsignificant results ($\chi^2[1, N = 339] = 1.07, p = .15$). However, the low sample size of 339 in the comparison would have required the $2.00 group to have a response rate more

Table 1. Cumulative Response Rates After Each Mailing

	Control and Experimental Conditions				
Mailing	Control	\$0.25	\$0.50	\$1.00	\$2.00
First***	54.2%	62.7%	63.1%	72.8%	77.6%
Second***	72.0	74.6	78.0	88.2	88.2
Third***	83.3	82.8	82.7	91.7	94.7
Fourth*	88.1	86.4	86.5	92.9	95.9
Sample Size	168	169	168	169	170

*** $p < .001$. ** $p < .01$. * $p < .05$, (χ^2, $df = 4$).

than 8.0% higher than the \$1.00 group to be judged significantly higher at $\alpha = .05$, whereas the sample response rate was only 4.8% higher.

SECOND MAILING

The cumulative response rate continued to be significantly different among the groups, following the second mailing, with the response rate trending upward with incentive amount (see Table 1). This is supported by a highly significant value of the likelihood ratio test statistic ($-2 \log\lambda = 25.61$, $p < .001$). To determine the incentive amounts that contributed to these differences, we first tested equality in the \$0.25 and \$0.50 groups and concluded that these groups are identical ($\chi^2[1, N = 337] = .54$, $p = .27$). The combined \$0.25–\$0.50 group compared against the control group yielded nonsignificant differences ($\chi^2[1, N = 505] = 1.07$, $p = .15$). Thus, while there is evidence that the \$0.25–\$0.50 group has a higher response rate than the control group, the difference in the sample, 76.3% − 72% = 4.3% is not significant. However, due to sample size, a difference of more than 7.0% would be required for significance at $\alpha = .05$.

Comparison of the \$0.25–\$0.50 group with the \$1.00 group indicated a significantly higher response rate in the \$1.00 group ($\chi^2[1, N = 506] = 10.06$, $p < .01$). It is obvious from Table 1 that no difference exists between the \$1.00 and \$2.00 groups.

THIRD AND FOURTH MAILINGS

In both mailings the likelihood ratio statistic was highly significant ($-2 \log\lambda = 21.62$, $p < .001$, for three waves, and $-2 \log\lambda = 13.89$, $p < .01$ for four waves), indicating monotonic increasing rates. However, it is clear from Table 1 that no significant difference in response rate exists between the control, \$0.25, and \$0.50 groups. Collapsing these groups into a single group and testing for a higher response rate in the \$1.00 group yields a highly significant difference after the third mailing ($\chi^2[1, N = 674] = 7.67, p < .01$), and a significant difference after the fourth mailing ($\chi^2[1, N = 674] = 3.93, p < .03$). In each case, the response rate from the \$2.00 group is 3% higher than that of the \$1.00 group so the \$2.00 group yields a significantly higher response rate than the 0–\$0.25–\$0.50 group as well. But a 5.4% difference would be required for a significant difference at $\alpha = .05$ between the \$1.00 and \$2.00 groups with the current sample sizes, therefore the differences are nonsignificant ($\chi^2[1, N = 339] = 1.20, p = .14$, for the third wave, and $\chi^2[1, N = 339] = 1.42, p = .12$, for the fourth wave).

DISCUSSION

Consistent with the majority of prior studies, it is clear that a \$0.25 incentive results in a significantly higher response rate than no incentive in a one-wave mailing. Although a \$0.50 incentive does not yield a significantly higher response rate than \$0.25 (regardless of the number of follow-up mailings), a \$1.00 incentive is effective in increasing the response rate significantly over the \$0.25 incentive. Although previous studies have concluded that the psychological effect of the inducement is what results in an increase in the response rates (e.g., Pressley and Tullar, 1977; Wotruba, 1966), this study shows that it is the actual monetary value of the inducement that plays a major role in increasing response rates. These results also show that this survey does not fit the model of Jobber and Saunders (1988) which claims that response rate is increased by a monetary incentive but that the amount of the incentive is immaterial.

What about the use of a combination of a monetary incentive and follow-up mailings? In the present study, the effect of the \$0.25 and \$0.50 incentives upon response rate are not significant after the second mailing. However, the \$1.00 or \$2.00 incentive results in a significantly higher response rate than the use of no incentive (or \$0.25 or \$0.50) when either one, two, three, or four mailings are employed. These results differ from Nederhof's study (1983) in which the incentive effect was diminished with follow-up mailings, but this is probably due to the more potent effect of monetary incentives, as well as to the larger incentives used in this study (Nederhof used a \$0.35 ballpoint pen). Thus, the question of whether or not an incentive effect "washes out" in the face of follow-up mailings depends upon the *value* of the incentive.

What implications do these results have for the survey researcher?

First of all, it should be noted that this study was performed with subjects who are highly educated (median = college education), and who have high annual family incomes (median = $51,276). Secondly, the topic of the survey, cable TV, is one of great interest to the public. These factors may account, in part, for the high response rate of 54% in the control group after the first mailing. However, in surveys with lower socioeconomic groups, and topics of less general interest, there would be more room for improvement percentagewise by using an incentive, and therefore it is possible that the monetary incentives would have an even greater effect. So, should the researcher decide, funds permitting, to include an incentive in the first mailing? The answer depends on the cost-benefit relationship and the quality of the data that one can expect from the incentive versus nonincentive options.

Cost-Benefit Analysis

RESULTS

The costs of labor, supplies, and mailings were calculated and divided by the cumulative number of returns for each wave and incentive condition, which provided a per unit cost for each questionnaire received (see Table 2). Predictably, the one-wave control condition incurred the lowest cost per return, but produced the lowest response rate. At the other extreme, the four-wave $2.00 incentive condition produced the highest response rate but also resulted in the highest total cost. The highest per-return cost was incurred by the one-wave $2.00 incentive, but this generated a 78% response rate. The per-return costs for the $0.25 and $0.50 conditions were comparatively high due to the additional postage required by the added weight of the coin(s) (39 cents versus 22 cents).

DISCUSSION

In a time-sensitive survey for which only one mailing is possible, our results suggest that an incentive of a quarter or a dollar should be used since the response rate increases significantly as the incentive amount increases from zero to a quarter to a dollar. The cost per return for each of these options is $3.54, $3.78, and $4.11 respectively (see Table 2). Although the $2.00 incentive produced the highest response rate from the first mailing in this study, recall that it was not significantly greater than that from the $1.00 incentive, and the cost per return of $5.15 is substantially larger.

Table 2. Cost Analysis of Experimental Conditions

Experimental Condition	Total Cost	Number Returned	Percentage Response	Cost per Return
1 Wave, $0.00	$321.33	91	54	$3.54
1 Wave, $0.25	$399.93	106	63	$3.78
1 Wave, $0.50	$442.82	106	63	$4.18
1 Wave, $1.00	$505.51	123	73	$4.11
1 Wave, $2.00	$679.34	132	78	$5.15
2 Wave, $0.00	$429.58	121	72	$3.56
2 Wave, $0.25	$493.13	126	75	$3.92
2 Wave, $0.50	$537.26	131	78	$4.11
2 Wave, $1.00	$591.37	149	88	$3.97
2 Wave, $2.00	$756.07	150	88	$5.05
3 Wave, $0.00	$497.62	140	83	$3.56
3 Wave, $0.25	$554.42	140	83	$3.97
3 Wave, $0.50	$590.85	139	83	$4.26
3 Wave, $1.00	$630.57	155	92	$4.07
3 Wave, $2.00	$794.52	161	95	$4.94
4 Wave, $0.00	$539.26	148	88	$3.65
4 Wave, $0.25	$595.69	146	86	$4.09
4 Wave, $0.50	$629.49	147	88	$4.29
4 Wave, $1.00	$657.49	157	93	$4.19
4 Wave, $2.00	$818.94	163	96	$5.03

NOTE: Incentives were included in the first-wave mailing only.

Another useful way to do cost-benefit analysis is to compute the cost-per-additional return as the incentive is increased. From "no-incentive" to a "quarter" the cost is $5.24 per additional return, whereas from "no-incentive" to a "dollar," the cost per additional return is $5.76. Since the increase in the response rate over the control condition is more than doubled by using the $1.00 versus the $0.25 incentive (19% versus 9%), when weighing both factors of cost and response rate, the $1.00 incentive would be the best choice in a one-wave mailing.

If a two-, three-, or four-wave survey is to be carried out, our results indicate that, considering the response rate, either a $1 incentive should be used, or none at all. However, although the $1 incentive always produces a significantly higher response rate than no incentive, from Table 1 it is clear that the increases in the response rate from the $1 incentive slow down considerably after the second mailing, once the response rate reaches 88.2%. This figure is virtually identical with the response rate obtained after four mailings in the no-incentive group. Thus, *as a practical matter,* there is no significant difference between the four-wave no-incentive group response rate and the two-wave $1 incentive group response rate. From Table 2, the cost per response for the four-wave no-incentive group is $3.65 per return, compared to $3.97 for the two-wave one-dollar incentive. Therefore if time permits, it would be more cost beneficial to conduct a four-wave survey without an incentive than a two-wave survey with a $1.00 incentive. However, the time needed to complete a four-wave survey also represents a cost to the organization conducting the survey, and it may be more beneficial to complete the survey within a briefer time frame with a two-wave survey and a dollar incentive at a total cost of less than 10% more.

Finally, the response rates after the third and fourth mailings without incentive were extremely good in the current survey (83% and 88%), and even though the same number of mailings with the one or two dollar incentives provide significantly higher response rates, the increase in *total* costs are high: about 25% for the $1 incentive and over 50% for the $2 incentive (see Table 2). Cost considerations suggest then, that three or four mailings without incentive would be more cost beneficial than the same number of mailings using the $1 or $2 incentive, provided that generally high response rates are predicted.

Response Bias

RESULTS

Demographic Differences. Demographic data on age (nine categories), education (six categories), and family income (ten categories) were analyzed for respondents following the *first* mailing to determine if there was intertreatment bias. As the variables are all ordinal, the Goodman-Kruskal gamma measure of association (Agresti, 1984) or the Kruskal-Wallis test (Blalock, 1979; Hays, 1981; Kruskal, 1952; Kruskal and Wallis, 1952; Siegel, 1956) can be applied. Since, a priori, one might expect respondent's level of education and income to decrease as the incentive amount increases, we tested for this using the Goodman-Kruskal gamma (γ) statistic. The results confirmed such a negative association ($\gamma[N = 555] = -.09$ and $\gamma[N = 527] = -.1$[illegible] education and income respectively, $p < .05$ in both cases), but the small values of γ, although significant, indicate the association is "mild." The Bonferroni method of simultaneously testing the ten pairs of groups (Miller, 1986) indicated that the only significant pairwise negative associations occurred between the $0.25 and $2.00 incentive groups for education ($\gamma[N = 237] = -.28, p < .005$) and between the control and \$2.00 incentive group for income ($\gamma[N = 214] = -.27, p < .005$). The Kruskal-Wallis statistic, corrected for tied ranks, indicated nonsignificant differences between respondent groups in age ($H[4, N = 555] = 1.61, p = .81$).

It is surprising that the only significant difference in educational level among pairs of incentive groups was found between the \$0.25 and \$2.00 groups, not between the no-incentive and \$2.00 groups as is the case with income. One possible explanation is that the \$0.25 group may have had a higher level of education than the other groups to begin with, despite our efforts to minimize such a possibility by using systematic selection and assignment methods. Since the response rates in these two groups were so high after four mailings, 86% and 96% respectively, we tested for differences after four mailings since the total respondents are likely to be nearly representative of the entire original groups. We found that a significant negative association still existed ($\gamma[N = 306] = -.17, p < .02$), leading us to conclude that the \$0.25 group as a whole probably had a higher level of education to begin with than the \$2.00 group.

In view of this result, it was prudent to test for differences in income level after four mailings between the no-incentive and \$2.00 groups, since this was the only pair that exhibited a significant negative association after one mailing. The response rates were 88% and 96% respectively in the control and \$2.00 groups after four mailings, so the respondents should be nearly representative of the groups as a whole. We found that a significant negative association still existed ($\gamma[N = 289] = -.13, p = .04$), leading us to conclude that the income level of the \$2.00 group was probably lower than that of the control group to begin with.

Fixed-Response Questions. The questionnaire contained 28 fixed-response (check-off) questions concerning subjects such as length of subscription, VCR ownership, movie rental, and interest in new services. Chi-square analyses indicated nonsignificant differences between the five respondent groups to the first mailing on answers to each of the 28 questions.

Respondent's Effort in Completing Questionnaires. The degree of effort made by respondents to the first mailing was measured by (*a*) the percentage of questions answered, (*b*) the number of distinct answers given to the short answer question, (*c*) the number of comments provided at the end of the questionnaire, and (*d*) the total number of words written in response to the short-answer and open-ended questions. The percentage of questions answered in the control and each of the experimental groups was nearly identical, ranging from 96% to 97%. However, there were differences in items *b*, *c*, and *d*, and we now discuss these in turn.

The short-answer question included four numbered lines in which respondents could write in as many as four short answers. Since we anticipated the number of short answers given (ranging from one to four) would increase with incentive amount, we tested for positive association with the Goodman-Kruskal gamma statistic, and the result was highly significant ($\gamma[N = 489] = .10$, $p < .01$). The mean number of short answers given increased steadily with incentive amount (1.53, 1.61, 1.72, 1.75, 1.81). Significant pairwise positive associations are indicated between the no-incentive and $1.00 groups ($\gamma[N = 194] = .18$, $p < .05$), and the no-incentive and $2.00 groups ($\gamma[N = 204] = .23$, $p < .01$).

The section at the end of the questionnaire asked for opinions, comments, or suggestions about the cable television company. We expected incentive amount to positively influence the number of comments provided (which we coded as: none, one, two, and three or more), and the Goodman-Kruskal gamma statistic indicated a positive association ($\gamma[N = 488] = .09$, $p < .05$). The mean number of comments in each group in order of increasing incentive amount is 0.58, 0.76, 0.72, 0.82, and 0.82. A significant pairwise positive association is indicated between the no-incentive and $1.00 (as well as $2.00) respondent groups ($\gamma[N = 194] = .203$, $p < .05$).

The median number of words written in order of increasing incentive amount is: 27.0, 31.5, 34.0, 37.0, and 38.0. These figures suggest that the number of words written increases with incentive amount. Since the distribution of the number of words is highly skewed to the right, we tested for a monotone increase using a nonparametric test based on the Kruskal-Wallis rank scores: the Jonckheere M-statistic (Miller, 1986:87–88). This statistic has an asymptotic normal distribution and the z-score is highly significant ($z[M] = 2.50$, $N = 488$, $p < .01$) indicating a monotonic relationship. Pairwise comparisons of the control group with each of the incentive groups using the Kruskal-Wallis one-tailed test indicate significant differences in all but the $0.25 group ($H[1, N = 169] = 1.09$, $p = .15$ for $0.25; $H[1, N = 170] = 3.82$, $p < .03$ for $0.50; $H[1, N = 194] = 4.31$, $p < .02$ for $1.00; and $H[1, N = 204] = 7.41$, $p < .005$ for $2.00).

Favorability of Comments. Comments made in the final section of the questionnaire were rated as to how favorable they were toward the cable television company: either generally positive (score of 2), neutral (1), or negative (0). Since favorability is an ordinal variable, we tested for a positive association with incentive amount using the gamma statistic, and the result was significant ($\gamma[N = 488] = .15$, $p < .02$). As the mean favorability score is 0.93 in the no-incentive group and between 1.02 and 1.07 in the other four groups, one might expect that the amount of incentive is immaterial in biasing the comments. This is indeed the case as the gamma statistic is not significant when testing the four incentive groups for differences ($\gamma[N = 405] = .05, p = .26$), and is highly significant when testing for differences between the incentive and non-incentive groups ($\gamma[N = 488] = .36, p < .01$). There are pairwise significant differences between the control group and: the \$0.50 group ($\gamma[N = 170] = .37, p < .02$); the \$1.00 group ($\gamma[N = 194] = .33, p < .02$); and the \$2.00 group ($\gamma[N = 204] = .43, p < .003$). Even testing the control group against the \$0.25 group produces a p-value of .06.

DISCUSSION

Prior research is equivocal on the effect of an incentive on the demographic distribution of respondents. In the Nederhof study (1983), the first survey resulted in a lower level of education for respondents in the incentive group ($N = 1{,}073$), but the second survey, conducted with smaller sample sizes (Ns = 667 and 99), resulted in nonsignificant differences. The inconsistent results may have been due to the lower statistical power of the second survey. The few studies that examined the effect of monetary incentives on the annual income of respondents did not find any differences (Dohrenwend, 1970; Erdos, 1957; Furse and Stewart, 1982). In the present study, although the statistics show that the educational and income levels of respondents to the first mailing tend to decrease slightly as the incentive amount increases, much, if not all, of this can be explained by the fact that the \$2.00 group most likely had a lower income and educational level than the other groups to begin with, and the \$0.25 group probably had the highest educational level. Thus, no strong conclusions can be made about the effect of the amount of incentive on education and income from this study.

The absence of differences in the answers to fixed-response questions is consistent with the results of previous studies (Goodstadt, Chung, Kronitz, and Cook, 1977; Hancock, 1940; Hansen, 1980; Heads and Thrift, 1966; Maloney, 1954; Mizes, Flees, and Roos, 1984; Nederhof, 1983). Factual information, and opinions about survey topics do not generally seem to be affected by incentives. The lack of differences in the percentage of questions answered as a function of monetary incentives is also consistent with most previous research (e.g., Furse and Stewart, 1982; Shuttleworth, 1931; Wotruba, 1966). Once individuals decide to fill out survey questionnaires, they appear to answer the same percentage of questions, regardless of monetary incentives.

The finding of a positive association between incentive amount and the number of short answers, comments, and words written is consistent with prior research that has found that monetary incentives tend to generate more comments, more completeness of response, and more clarifying information (Godwin, 1979; Shuttleworth, 1931).

Previous research on favorability of comments as a function of monetary incentives is equivocal. Trice (1984) found that hotel guests offered a $1 tariff reduction in a hotel fee gave fewer negative ratings about the hotel than those not offered the incentive. However, a survey of magazine readers failed to find differences in the value of a magazine on a five-point scale as a function of receiving a $0.25 incentive (Goodstadt, Chung, Kronitz, and Cook, 1977). The present study found that the comments made were more favorable toward the survey sponsor when respondents were paid an incentive of any amount.

Conclusion

The results of this study, consistent with previous research, indicate that a $0.25 incentive significantly increases the response rate over a no-incentive control condition in a one-wave survey. This study also shows, however, that a larger incentive of $1.00 or $2.00 significantly increases the response rate over that obtained with a $0.25 incentive. This finding suggests that it is not solely the psychological effect of an inducement that results in an increased response rate, but that the actual monetary value also plays an important role.

When comparing the use of an incentive in a one-wave survey versus a multiple wave survey without incentives, the results indicate that a combination of techniques is more effective than either technique used alone. A combination of a $1 or $2 incentive enclosed with the first mailing and zero, one, two, or three follow-up mailings results in a significantly higher response rate than an equivalent number of mailings sent without an incentive.

A cost-benefit analysis indicates that enclosure of a $1 incentive with the first mailing in a one-, two-, three-, or four-wave survey is an effective and cost-beneficial method of increasing response rates. However, the response rate achieved with four mailings without incentives was also very high—88% (compared to 93% with a $1 incentive, and 96% with a $2 incentive, both in a four-wave survey). Therefore, a well-designed Dillman-type survey can produce satisfactorily high response rates without the use of incentives, especially if the topic is of great interest to the public and the subjects are highly educated.

Comparison of the responses provided by respondents in the incentive and no-incentive conditions revealed evidence of intertreatment bias. A positive association was found between the incentive amount and the degree of effort expended, as measured by the number of words, comments, and short answers provided by the respondents.

Those who received incentives also made comments that were more favorable to the survey sponsor. The extra respondent effort produced by the monetary incentives may be desirable in some surveys, particularly when open-ended questions are asked, or when respondents are asked to check their records for information. But the finding that more favorable comments are made by rewarded subjects should be kept in mind prior to using monetary incentives, particularly when critical responses are needed.

References

Agresti, A. (1984)
Analysis of Ordinal Categorical Data. New York: John Wiley.
Armstrong, J. S. (1975)
"Monetary incentives in mail surveys." Public Opinion Quarterly 39:111–116.
Blalock, H. M. (1979)
Social Statistics. New York: McGraw-Hill.
Blumberg, H. H., C. Fuller, and A. P. Hare (1974)
"Response rates in postal surveys." Public Opinion Quarterly 38:113–123.
Dillman, D. A. (1978)
Mail and Telephone Surveys. New York: John Wiley.
Dohrenwend, B. S. (1970)
"An experimental study of payments to respondents." Public Opinion Quarterly 34:621–624.
Erdos, P. L. (1957)
"Successful mail surveys: High returns and how to get them." Printer's Ink 258:56–60.
Furse, D. H., and D. W. Stewart (1982)
"Monetary incentives versus promised contribution to charity: New evidence on mail survey response." Journal of Marketing Research 19:375–380.
Fox, R. J., M. R. Crask, and J. Kim (1988)
"Mail survey response rate: A meta-analysis of selected techniques for inducing response." Public Opinion Quarterly 52:467–491.
Godwin, R. K. (1979)
"The consequences of large monetary incentives in mail surveys of elites." Public Opinion Quarterly 43:378–387.
Goodstadt, M. S., L. Chung, R. Kronitz, and G. Cook (1977)
"Mail survey response rates: Their manipulation and impact." Journal of Marketing Research 14:391–395.
Hancock, J. W. (1940)
"An experimental study of four methods of measuring unit costs of obtaining attitude toward the retail store." Journal of Applied Psychology 24:213–230.
Hansen, R. A. (1980)
"A self-perception interpretation of the effect of monetary and nonmonetary incentives on mail survey respondent behavior." Journal of Marketing Research 17:77–83.
Hays, W. L. (1981)
Statistics. New York: Holt, Rinehart and Winston.
Heads, J., and H. Thrift (1966)
"Notes on a study in postal response rates." Commentary 8:257–262.
Heberlein, T. A., and R. Baumgartner (1978)
"Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature." American Sociological Review 43:447–462.

Houston, M. J., and N. M. Ford (1976)
"Broadening the scope of methodological research on mail surveys." Journal of Marketing Research 13:397–403.
Jobber, D., and J. Saunders (1988)
"Modeling the effects of prepaid monetary incentives on mail-survey response." Journal of the Operational Research Society 39:365–372.
Kanuk, L., and C. Berenson (1975)
"Mail surveys and response rates: A literature review." Journal of Marketing Research 12:440–453.
Kephart, W. M., and M. Bressler (1958)
"Increasing the responses to mail questionnaires: A research study." Public Opinion Quarterly 22:123–132.
Kruskal, W. H. (1952)
"A nonparametric test for the several sample problem." Annals of Mathematical Statistics 23:525–539.
Kruskal, W. H., and W. A. Wallis (1952)
"Use of ranks in one-criterion variance analysis." Journal of the American Statistical Association 47(260): 584–621.
Linksy, A. S. (1975)
"Stimulating responses to mailed questionnaires: A review." Public Opinion Quarterly 39:82–101.
Maloney, P. W. (1954)
"Comparability of personal attitude scale administration with mail administration with and without incentive." Journal of Applied Psychology 38:238–239.
McDaniel, S. W., and C. P. Rao (1980)
"The effect of monetary inducement on mailed questionnaire response quality." Journal of Marketing Research 17:265–268.
Miller, R. G., Jr. (1986)
Beyond ANOVA, Basics of Applied Statistics. New York: John Wiley.
Mizes, J. S., E. L. Fleece, and C. Roos (1984)
"Incentives for increasing return rates: Magnitude levels, response bias, amid format." Public Opinion Quarterly 48:794–800.
Nederhof, A. J. (1983)
"The effects of material incentives in mail surveys: Two studies." Public Opinion Quarterly 47:107–111.
Pressley, M. M., and W. L. Tullar (1977)
"A factor interactive investigation of mail survey response rates from a commercial population." Journal of Marketing Research 14:108–111.
Robertson, D. H., and D. N. Bellenger (1978)
"A new method of increasing mail survey responses: Contributions to charity." Journal of Marketing Research 15:632–633.
Robertson, T., F. W. Wright, and R. L. Dykstra (1988)
Order Restricted Statistical Inference. New York: John Wiley.
Robinson, R. A., and P. Agisim (1951)
"Making mail surveys more reliable." Journal of Marketing 15:415–424.
Shuttleworth, F. K. (1931)
"A study of questionnaire technique." Journal of Educational Psychology 22:652–658.
Siegel, S. (1956)
Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.
Trice, A. D. (1984)
"Hotel ratings: IV. Recruitment and incentives." Psychological Reports 55:58.
Wotruba, T. R. (1966)
"Monetary inducements and mail questionnaire response." Journal of Marketing Research 3:393–400.
Yu, J., and H. Cooper (1983)
"A quantitative review of research design effects on response rates to questionnaires." Journal of Marketing Research 20:36–44.

EFFECTS OF INTERVIEW MODE ON SELF-REPORTED DRUG USE

WILLIAM S. AQUILINO AND
LEONARD A. LO SCIUTO

Abstract The effects of interview mode on 18- to 34-year-olds' self-reported tobacco, alcohol, marijuana, and cocaine use were investigated. RDD telephone and personal drug use surveys of the state of New Jersey were conducted in 1986–87. In the personal interview, respondents recorded their drug use on self-administered answer sheets. Compared to the area probability sample in the personal interview, RDD yielded a sample of blacks higher in income and education, and more likely to be married and currently employed; white SES was also higher in the telephone sample, but to a lesser degree than for blacks. Controlling for demographic characteristics and RDD's exclusion of nontelephone households, the telephone survey yielded significantly lower estimates of blacks' alcohol consumption, and lifetime and recent marijuana use. Whites' alcohol consumption was slightly lower by telephone; otherwise, estimates of whites' use of the four substances were nearly identical in the two modes. Sample coverage, respondent demographic characteristics, and racial matching of interviewer and respondent did not account for the significant mode differences. Characteristics of the interview situation itself, such as provision of privacy in the self-administered format, may have influenced tendencies toward socially desirable responding to a threatening topic such as drug use.

Telephone interviewing has increased in popularity relative to personal interviewing due to cost savings, efficiency, and the availability of com-

WILLIAM S. AQUILINO is a Research Scientist with the Center for Demography and Ecology, Department of Sociology, University of Wisconsin—Madison. LEONARD A. LO SCIUTO is Director, Institute for Survey Research, Temple University. The study was supported by grant No. R01-DA04280 from the National Institute on Drug Abuse, Bethesda, Maryland, and by PHS National Research Service Award #AG00129-03 from the National Institute on Aging. Surveys were conducted by the Institute for Survey Research at Temple University. The personal interview data were collected by the New Jersey Household Survey on Drug Abuse, funded by the Department of Public Health, State of New Jersey. The authors wish to thank Linda Thompson for helpful comments on an earlier version of this paper.

Public Opinion Quarterly Volume 54:362–395 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

puter-assisted methods to speed data collection and processing (Groves, 1983; Klecka and Tuchfarber, 1978). Compared to face-to-face interviewing, telephone surveys are cheaper, have shorter data collection periods, and lend themselves to closer monitoring of interviewer performance, thus enhancing quality control during data collection.

The utility of the telephone survey has been demonstrated in marketing, political polling, health, and numerous forms of attitudinal research. Herzog and Rodgers (1988), for example, found almost no differences in response distributions obtained from telephone and personal interviews in a survey of older Americans' political attitudes, health, well-being, and life satisfaction. Telephone approaches have been shown to produce reliable and valid results for relatively nonthreatening interview topics. Much less is known, however, about the efficacy of telephone surveys for highly threatening interview topics, such as sexual behavior, criminality, and the use of illicit drugs. Surveys on these topics have relied almost exclusively on costly and time-consuming face-to-face interviewing. This research addressed the question: To what extent does the validity of sensitive surveys differ by mode of interview? Specifically, is there a difference between face-to-face surveys using self-administered questionnaires (SAQs) and telephone surveys in their estimates of lifetime and current use of tobacco, alcohol, marijuana, and cocaine? Our research questions centered on comparing telephone and personal survey modes in terms of nonresponse, demographic characteristics of the respondents, and validity of drug use estimates: (1) what differences in sample demographic characteristics are produced by multistage area probability sampling versus Random Digit Dial (RDD) methods; and (2) do personal surveys using SAQs and telephone surveys provide different estimates of drug use for the same population? If so, to what extent are mode differences attributable to sample coverage, nonresponse bias, or to the interview situation itself?

Literature Review

RESPONSE RATES

Response rates typically vary between telephone and face-to-face interview modes. Experimental comparison of interview modes for the National Health Interview Survey (NHIS) yielded overall response rates of 80% for the telephone, and 96% for the personal survey (Groves, Miller, and Cannell, 1987). Groves and Kahn (1979) reported an RDD response rate of 59% in their national survey (including all unanswered telephone numbers in the denominator of the response rate equation), compared to 75% in the personal interview. Exclusion of the unanswered numbers raised the telephone rate to 70%.

Telephone surveys tend to have higher break-off rates (Groves and Kahn, 1979) and higher refusal rates (Jordan, Marcus, and Reeder, 1980) than personal surveys. The collection of household composition data early in the call, to determine eligibility and respondent selection, may increase refusals (Hauck and Cox, 1974). Telephone nonresponse occurs disproportionately among minorities, the less affluent, less educated, and among older respondents (Freeman et al., 1982; Groves, Miller, and Cannell, 1987; O'Neil, 1979). Nonresponse may vary by race and ethnicity. Blacks tend to be disproportionately inaccessible in telephone surveys (Weaver, Holmes, and Glenn, 1975); this is especially true for younger, lower income, and male black respondents.

SAMPLE COVERAGE

The survey literature is clear in suggesting that RDD sampling produces a different demographic profile of respondents than multistage area probability sampling. The sampling frame for RDD excludes households without telephones, which account for approximately 7% of households in the United States (Thornberry and Massey, 1988). The demographic correlates of nontelephone status make this a potentially serious bias for drug use surveys. Respondents from nontelephone households are more likely than those with telephones to be nonwhite, of lower income and education, and to be never married, divorced, or separated (Thornberry and Massey, 1988; Tull and Albaum, 1977). RDD's exclusion of nontelephone households also underrepresents rural nonfarm areas, single heads of household, persons under age 35, divorcees, and renters (Freeman et al., 1982). Income is the strongest correlate of nontelephone status; nearly 30% of families with incomes below $5,000 in 1986 were without telephone service (Thornberry and Massey, 1988). Klecka and Tuchfarber (1978) compared RDD and face-to-face samples in a Cincinnati survey of crime victimization. The RDD sample captured fewer poor households and had a significantly higher education level than the personal survey. This literature suggests that, compared to the area probability sample in the personal survey, the RDD sample will have significantly higher income and educational attainment, and will have proportionately fewer nonwhite, fewer divorced, and fewer separated respondents.

BIAS IN DRUG USE ESTIMATES

Drug use surveys are subject to response set bias, most notably socially desirable responding in the face of threatening survey items

(Mensch and Kandel, 1988). Use of illicit drugs, such as marijuana and cocaine, and heavy versus social drinking, are assumed to be socially undesirable behaviors for the majority of respondents. To the extent that social desirability affects responses, and respondents feel threatened by such interview questions, one would expect *underreporting* of drug use to be the largest threat to the validity of drug use surveys (Bradburn et al., 1978). The more socially unacceptable the substance, the greater the underreporting should be. Sudman and Bradburn (1974) maintained that the higher the threat, the more influential the interview milieu on respondents. There is reason to suspect, then, that the more threatening the topic, the more potential there is for survey characteristics, such as interview mode, to affect tendencies toward socially desirable responding. The central question of this paper, then, is: Will a switch from face-to-face to telephone survey modes affect the validity of population estimates of drug use, i.e., result in lower estimates of drug use?

A recent study by Mensch and Kandel (1988) demonstrated the susceptibility of drug use reports to socially desirable responding. Underreporting of lifetime marijuana use in the National Longitudinal Survey of Youth (NLSY) was related to familiarity with the interviewer, who may have conducted up to three interviews with the same respondent. Panel members reinterviewed by the same person were more likely than those with different interviewers to deny marijuana and cocaine use on subsequent interviews, even after admitting use at the initial interview. Familiarity with the interviewer increased respondents' propensity toward socially desirable responding.

There has been relatively little research on willingness to report highly threatening or undesirable behaviors in telephone versus face-to-face interviews. Several researchers have expressed skepticism about the ability of telephone surveys to secure honest reporting of sensitive or illicit behavior (Freeman et al., 1982; Johnston and O'Malley, 1985; Sudman, 1976). Groves (1979) found that respondents expressed greater uneasiness about discussing sensitive topics on the telephone than in face-to-face interviewing. Telephone interviewing may elicit higher refusal rates for sensitive questions than in-person interviewing (Groves and Kahn, 1979).

Studies comparing personal and telephone surveys for response set bias have produced mixed results (Orwin and Boruch, 1982). The validity of telephone results has been demonstrated with relatively nonthreatening topics: the proportion of women admitting ever drinking alcohol (Hochstim, 1967); physicians' self-reported rates of reading medical journals (Colombotos, 1969); voting (Rogers, 1976); Catholics' attitudes toward abortion and birth control (Wiseman, 1972); and health-related surveys (Groves, Miller, and Cannell, 1987; Herzog et al., 1983; Siemiatycki, 1979). Although these studies support the validity of telephone interviewing, their results are based on less threatening topics and may not be generalizable to highly threatening surveys concerned with self-reports of illicit or embarrassing behaviors.

Comparisons involving more threatening surveys suggest that telephone interviewing may be more susceptible than personal surveys to social desirability effects. Respondents (randomly assigned to interview mode) were more likely to report psychiatric symptoms and depression in face-to-face interviews than by telephone (Henson, Cannell, and Roth, 1978); face-to-face respondents in this study also had significantly lower social desirability scores on the Crowne-Marlowe scales than telephone respondents. Herman (1977), in a mixed-mode survey of employees' voting for unionization, found telephone respondents less willing than those personally interviewed to reveal sensitive information such as unlawful campaign practices and how they intended to vote. Respondents were more reluctant in telephone than in personal interviews to reveal financial information, especially their indebtedness, and current and projected income (Schmiedeskamp, 1962).

One of the best methods for reducing response set effects due to social desirability may be the use of self-administered questionnaires (SAQs) in the context of a personal interview (Sudman and Bradburn, 1974). Unfortunately, the SAQ has been all but ignored in studies comparing interview modes (apart from Hochstim's [1967] comparison of mail, telephone, and personal modes). With SAQs, interviewers read instructions, and possibly the questions, but do not see respondents' answers. The National Household Survey on Drug Abuse (NHSDA) and the personal New Jersey drug survey reported here, used SAQs to maximize respondents' privacy in reporting drug use in the personal interview. Use of SAQs should increase the likelihood of finding differences between face-to-face and telephone surveys in responses to sensitive questions.

To summarize, direct comparisons of survey modes have not provided unequivocal evidence concerning the ability of telephone versus personal surveys to elicit accurate reports of socially undesirable behavior. However, most comparisons of interview modes have relied on topics of relatively mild threat compared to drug use surveys. The survey literature suggests that characteristics of the interview situation become increasingly important in affecting responses as survey items become more threatening. To the extent that questions concerning heavy drinking, marijuana use, and cocaine use are highly threatening, telephone interviews can be expected to be more susceptible to social desirability bias than face-to-face interviews with SAQs for threatening questions. Thus telephone surveys should yield lower estimates of drug use than the personal surveys, and the degree of underreporting should vary with the sensitivity of the question. Telephone estimates of heavy drinking, recent marijuana use, and recent cocaine use should be more susceptible to underreporting than estimates of social drinking or tobacco use.

Study Design

The objective of state and national drug use surveys is to estimate the incidence, prevalence, and correlates of the use and abuse of alcohol and other drugs in a representative sample of the household population. The NHSDA, for example, is conducted biannually to monitor national trends in drug use. Recently, several state governments have undertaken surveys to chart statewide drug use trends and to use as a resource for allocating drug abuse prevention funds. One such survey, conducted in 1986–87 for the state of New Jersey, provided the face-to-face interview data for this study.

DESIGN OF THE PERSONAL SURVEY

Questionnaire Design. The New Jersey face-to-face survey used the identical survey instruments developed for the 1985 NHSDA, and replicated the data collection procedures of the NHSDA exactly. Drug categories included: tobacco, alcohol, marijuana, cocaine (crack was added to the state survey), opiates, hallucinogens, and the nonmedical use of prescription drugs. Additional questions included health and behavioral consequences of drug use, and demographic characteristics of the respondent. The personal interview lasted 45 to 60 minutes, on the average.

Anonymity. For all but tobacco, answers to drug use questions were recorded by respondents on self-administered answer sheets; interviewers read the instructions at the start of each drug sequence, and, at the respondent's request, read the questions aloud, while the respondent completed the answer sheet. Respondents were told that interviewers would not see their answers on the self-administered sheets. Answer sheets were sealed in an envelope in the respondent's presence, upon completion of the interview. No names were recorded on the questionnaires, answer sheets, or envelopes containing the answer sheets. Respondents were offered the option of mailing the completed questionnaire directly upon completion of the interview (accompanying the interviewer to the nearest mailbox), although few exercised this option.

Data Collection. A sample of New Jersey's civilian, noninstitutionalized household population aged 18–34 years was developed. The multistage area probability design yielded a total sample of 4,571 residential addresses.[1] Advance letters were sent to selected households in the sample. The letter explained the purpose of the survey, confidentiality procedures, and the voluntary nature of participation. All screening and interviewing was done in-person. During screening, household members were listed by age and first name, and one respondent per household was selected from the eligible 18- to 34-year-olds using tables developed by Kish (1965). When definite refusals were obtained, one refusal conversion attempt was made, with reassignment to another interviewer where possible. A screening rate of 90.5% was achieved, and interviews were completed in 79.1% of the households determined to include an eligible respondent aged 18 to 34 (N = 1,042 completed interviews). The field period extended from June, 1986, through January, 1987.

DESIGN OF THE TELEPHONE SURVEY

The overriding concern in designing the telephone survey was comparability to the personal survey. Thus, it was paramount that the telephone survey attempt to (1) achieve high screening and interview response rates, (2) minimize item nonresponse, (3) preserve question content and meaning, and (4) guarantee confidentiality to respondents.

In order to maximize response rates, we decided to conduct a telephone survey averaging no more than 25 minutes. It was clear that only a subset of the personal interview sequences could be used, since administering the entire NHSDA questionnaire over the phone would have required well over an hour. We decided to ask the complete sequence of items about a smaller number of drugs: tobacco, alcohol, marijuana, and cocaine (including crack). These drugs were selected because they are relatively prevalent in the population, they have potential for abuse, they represent both legal and illegal substances, and they form a continuum from relatively nonthreatening (smoking, social drinking) to threatening (heavy drinking, marijuana use, cocaine use) questions.

Questionnaire Revision and Pretesting. This phase of the research involved reformatting and pretesting the NHSDA drug use and demographic items for telephone administration. Four waves of pretests (a total of 75 interviews) were conducted in August and September, 1987. The telephone numbers for the pretest interviewing were randomly selected from the telephone directory. Between each wave, the questionnaire was revised based on interviewer comments and the results of the interviews. Two types of changes were necessary to ensure the success of the telephone survey: reformatting items with many precoded response categories and the introduction of branching patterns within the drug sequences.

---

[1] A technical report with detailed descriptions of sampling procedures and response rate calculations is available upon request from the first author.

In the personal interview, items asking frequency of drug use presented respondents with up to 8 response categories, while income had 12 categories presented in a show card. Remembering the distinctions among 8 to 12 categories represents a far greater burden for the telephone than the personal respondent (Miller, 1984). To reduce respondent burden, these items were broken into subitems that recreated the original NHSDA precoded scales. Research has shown that the two-step (Miller, 1984) and bracketing (Locander and Burton, 1976) procedures for asking questions with many response categories produce results comparable to the single item format. In the revision for telephone, stem question wording and all original response categories were preserved.

In the face-to-face survey, respondents answered all questions within a drug sequence, regardless of whether or not they had used the drug, or how recently they used it. Thus, respondents were asked how often they had used marijuana in the last 30 days, even if previously they had indicated no use of marijuana in their life. This was done to prevent respondents from denying drug use merely to shorten interview time. This structure was impossible to administer over the telephone. Early rounds of pretests clearly indicated that redundancy in item sequences was unacceptable to respondents, and generated high degrees of irritation and interview break-offs. For the telephone interview to succeed, then, it was essential to introduce branching patterns in the telephone drug sequences.

It is important to note that all changes to the NHSDA were essential to minimize the telephone survey's interview and item nonresponse rates, respondent irritation, and interview break-offs. Failure to make these changes would have increased nonresponse in the telephone survey and, thus, jeopardized the comparison to the personal mode.

Data Collection. The Waksberg procedure (Waksberg, 1978) of Random Digit Dialing (RDD) was used to draw the telephone sample. The Waksberg procedure was adopted to maximize the proportion of residential households contacted during screening, thus reducing survey costs and field period length. Our goal was the completion of at least 2,000 interviews with respondents 18- to 65-years-old. (This paper reports substantive results for 18- to 34-year-olds only, the same age group available in the face-to-face survey.)

All residential households with at least one member 18- to 65-years-old were eligible for the telephone sample. Interviewers verified the number dialed, asked the screening informant if the number was residential, and, if so, obtained a listing by first name and age of all eligible household members. One respondent was randomly chosen for the interview using selection tables developed by Kish (1965). Screening informants were told that the interview concerned health-related issues; selected respondents were informed that the survey concerned tobacco, alcohol, and drug use. All initial refusers were recontacted by a different interviewer for refusal conversion. The data collection period for the telephone survey was October through December, 1987.

The Waksberg procedure produced a sample of 6,932 telephone numbers; 2,075 interviews were completed with adults 18- to 65-years-old, of which 864 were with 18- to 34-year-old respondents. The screening response rate was 71.5%. The interview response rate, the proportion of all selected respondents who completed the entire interview, was 82.5% for 18- to 34-year-old respondents. The difference in completion rates between the telephone and personal surveys was entirely due to differential screening nonresponse; the interview completion rate with selected respondents was slightly higher by telephone (82.5%) than face-to-face (79.1%, difference not statistically significant).

Results

MODE DIFFERENCES IN SAMPLE CHARACTERISTICS

One source of demographic differences between RDD and area probability sampling is RDD's exclusion of nontelephone households. Table 1 shows household telephone status by demographic characteristics for the face-to-face sample (the number of telephone lines in the household was asked in the personal as well as telephone interview). Blacks and other minorities were much more likely than whites not to have telephones. Respondents without telephones were also much more likely to be unemployed, have less than a high school education, and have lower incomes. Telephone status is correlated with respondents' race and economic status, both important predictors of drug use.

To estimate mode differences in demographic profiles, all nontelephone households were dropped from the face-to-face sample. Over 13% (21 cases) of black respondents, and 17% of the Hispanic cases were dropped from the face-to-face sample. For whites, very few households were without phone service (13 cases, or 1.7% of the total).

Demographic characteristics of the telephone and face-to-face samples are presented in Table 2. The percentage distributions show substantial differences for blacks and fewer differences for whites and Hispanics. For blacks, the RDD sample resulted in significantly differ-

Table 1. Telephone Status by Sample Characteristics, Personal Survey

	N of Cases	Percentage with No Phone
Total sample	1,041	5
Age		
18–21	222	5
22–25	223	5
26–29	249	6
30–34	339	4
Sex		
Male	433	4
Female	608	6
Race/ethnicity*		
Black	160	13
White	752	2
Hispanic	129	15
Marital status		
Married	430	4
Sep./div./wid.	89	9
Never married	521	5
Education*		
Less than high school	113	20
High school graduate	395	4
Some college	286	3
College graduate	247	2
Work situation*		
Employed full-time	675	4
Employed part-time	120	2
Unemployed	71	16
Not in labor force	175	7
Student status		
Full-time student	128	4
Part-time student	85	1
Nonstudent	824	6
Income*		
Under $7,000	358	8
$7,000 to $14,999	219	6
$15,000 to $29,999	304	2
$30,000 or more	137	2

NOTE: *N*s sum to less than 1,042 due to missing values.

*$p < .01$ for contingency table chi-square test of statistical independence (presence of telephone × sample characteristic).

Table 2. Respondent Characteristics by Race and Interview Mode, Nontelephone Households Excluded from Personal Sample (Percentage Distributions)

	Black		White		Hispanic	
	Telephone	Personal	Telephone	Personal	Telephone	Personal
N of cases	104	137	636	734	60	82
Age						
18–21	19	23	18	22	25	21
22–25	18	21	22	21	20	21
26–29	26	26	24	24	25	19
30–34	37	29	37	33	30	40
Sex						
Male	34	34	45	44	33	36
Female	66	66	55	56	67	64
Marital status						
Married	27	17*	42	43	39	49
Widowed	2	1	—	—	—	1
Divorced/separated	5	15	7	6	15	8
Cohabiting	7	6	5	2	3	1
Never married	60	62	46	48	40	41

Education						
Less than high school	10	14	5	7	18	25*
High school graduate	42	45	37	38	43	40
Some college	36	27	26	29	15	27
College graduate	12	14	32	26	23	9
Work situation						
Employed full-time	78	59*	72	67	57	57
Employed part-time	10	9	13	13	17	10
Unemployed	4	12	4	5	5	10
Not in labor force	9	19	11	15	22	24
Student status						
Full-time student	11	10	10	13	17	12
Part-time student	10	7	11	10	10	4
Nonstudent	80	84	79	77	73	84
Income						
Under $7,000	16	40*	20	32*	30	47
$7,000 to $14,999	24	21	16	21	33	21
$15,000 to $29,999	50	32	40	32	23	24
$30,000 or more	11	7	24	15	13	8

* $p < .05$ for contingency table chi-square test of statistical independence (sample characteristic × mode of interview).

ent distributions for marital status, employment, and income. The RDD sample yielded greater proportions of married (27% to 17%) and lower proportions of divorced or separated respondents (5% to 15%) than the personal interview sample. There were significantly lower numbers of unemployed blacks (4% to 12%), and higher numbers of full-time employed (78% to 59%), in the telephone sample. The largest difference for blacks was in income: by telephone, only 16% of blacks reported personal incomes below $7,000, compared to 40% for the in-person survey. By telephone, 61% of blacks reported incomes of $15,000 or more, compared to only 39% in the face-to-face interview. For the 18- to 34-year-old population, then, the switch from area probability to RDD sampling shifted the SES profile of black respondents substantially upward.

Although not statistically significant, the education distribution for blacks also is consistent with RDD's higher SES profile for blacks. The telephone sample produced fewer black respondents with less than a high school education and more who had completed at least some college.

The higher SES profile in the RDD sample was also evident for Hispanics, although not to the same degree as for blacks. The RDD sample yielded significantly more high school and college graduates than did the face-to-face sample, fewer Hispanic respondents with incomes below $7,000 (30% to 47%), and fewer unemployed respondents (5% to 10%).

The demographic profiles of the two survey modes were most similar for whites. Although RDD also furnished significantly fewer low income respondents than the area probability sample, the magnitude of the difference for whites was smaller than for blacks or Hispanics. The RDD sample identified slightly more cohabiting respondents than did face-to-face. Otherwise, the demographic profiles for whites in the two surveys were nearly identical.

Controlling for telephone status did not equalize the demographic profiles of the RDD and area probability samples. It is very likely, then, that the differences are attributable to differential nonresponse by race/ethnicity in RDD versus area probability samples. That is, after correcting for bias introduced by sampling frame differences, the racial/ethnic patterns of screening and interview nonresponse may differ between the two survey modes. Even if proportional numbers of lower SES minority households were drawn into the RDD and area probability samples, telephone screening and interviewing appear to have produced higher levels of nonresponse (refusals, noncompleted interviews) for low SES households than did face-to-face methods.

It should be noted that the black and Hispanic samples were two-thirds female in both surveys. The low proportion of minority males in the surveys may reduce the descriptive accuracy of the black and Hispanic demographic profiles in both modes.

MODE DIFFERENCES IN SELF-REPORTED DRUG USE

Four substances of high use-prevalence were chosen for these analyses: tobacco (cigarette smoking), alcohol, marijuana, and cocaine. Verbatim questions and response categories for drug use items are listed in the Appendix. In analyzing drug data, responses from both surveys were pooled into one data set, with mode of interview added to the pooled data set as a categorical variable. Demographic characteristics were entered as control variables in the analyses. Race-of-interviewer effects were also explored, but were found to be unrelated to the interview mode and respondent race effects described below. Results concerning the racial matching of interviewer and respondent (Tables 5 and 6) are presented later in this section, after findings for interview mode and respondent racial differentials have been described.

Comparisons of drug use estimates will be reported for black and white respondents only. The telephone survey obtained only 60 interviews with Hispanics in the 18- to 34-year-old age range, a sample size too small to allow for adequate demographic controls and tests for interactions in the multiple classification analyses. The results described below are based on telephone interviews with 104 blacks and 636 whites, and personal interviews with 160 blacks and 747 whites.

Dependent Variables. Our original intent in selecting dependent variables was to include, for each drug category, measures of lifetime prevalence, recent use (within the past 12 months), and current use (within the past 30 days). This was done for marijuana and cocaine only. For tobacco, lifetime prevalence (having tried a cigarette even once) seemed fairly trivial, and was replaced with the proportion who had smoked five packs or more in their lives. For alcohol, both the lifetime and past year prevalence had to be dropped due to lack of variance; 98% of whites, for example, had used alcohol at least once in their lifetimes.

In addition to prevalence and incidence estimates, variables were selected to show the amount of drug use, both current (e.g., the number of cigarettes smoked per day, the number of days a person had a drink, smoked marijuana, or used cocaine over the last 30 days), and lifetime (total number of times respondents used marijuana or cocaine in their lives). Variables were selected to furnish a representative picture of lifetime and current usage of the four drugs selected for the study.

Data Analyses. To control for sample coverage, respondents from nontelephone households were dropped from the personal survey. Analyses are based on weighted drug use estimates. Case weights reflected adjustments for household selection probabilities (including number of telephone lines for the telephone sample), for differential nonresponse, and for Census profiles of New Jersey's age, race, and sex composition. Analysis of unweighted data yielded no important differences from the weighted results reported below.

Multiple classification analysis was used to estimate net differences by interview mode and by race, and the significance of the mode and race effects. Sample demographic characteristics were entered as control variables in the models. The *net differences* reported in Table 3 (telephone minus personal mode estimates) and Table 4 (black minus white estimates) control for sex, age, employment status, income, education, student status, and marital status. Preliminary analyses indicated significant race by mode of interview interactions for both alcohol and marijuana. Therefore the decision was made to analyze mode effects for blacks and whites separately. It should be noted that the term "mode effects" in this study refers (for all drugs but tobacco) to the comparison of telephone with a *self-administered format* in the personal mode. The results are not generalizable to personal surveys with interviewer-administered drug sequences.

Tobacco. Mode effects by race of respondent are presented in Table 3. Cigarette smoking is the least sensitive or threatening survey topic of the four substances, and, following Sudman and Bradburn (1974), should have been the least susceptible to interview mode effects. This prediction is borne out by the data. There were no significant mode effects for blacks or whites on the five measures of smoking. Overall, then, the results suggest that telephone estimates of smoking, when corrected for demographic characteristics of the sample, are comparable to face-to-face estimates for both black and white populations.

Alcohol. Questions concerning recent drinking and heavy drinking are more sensitive than smoking, given the negative social reactions to drunkenness and alcoholism. These questions should be more reactive to interview mode effects. Consistent with this prediction, there were substantial mode effects for alcohol use, after controlling for demographic characteristics of the samples (see Table 3). For blacks, the telephone survey yielded lower estimates for all five drinking measures. The results are particularly striking for blacks' drinking over the 30 days prior to the interview. The telephone interview estimates for the number of drinking days, total drinks for the month, and the number of days respondents had five or more drinks were significantly lower than in the personal interview.

For whites, the proportion of respondents who drank in the last 30 days and the proportion who admitted getting drunk at least once in the past 12 months were significantly lower in the telephone interview, although the magnitude of the differences was small (about 6 percentage points in each case). The difference for blacks who drank in the last month was slightly larger (8 percentage points), but was not significant due to the smaller sample size. On the whole, the mode differences were much stronger for blacks' than for whites' drinking.

Marijuana. The race by mode interaction is most evident in reports of marijuana use. For blacks, the marijuana results were very much the same as for alcohol. The telephone survey furnished consistently lower estimates of blacks' lifetime and recent marijuana use than the face-to-face survey. The number of current users was over 50% lower in the telephone than in the personal interview, while the number of lifetime users was about 25% lower in the telephone survey. For blacks, the telephone mode also furnished significantly lower estimates of both the days and the amount of marijuana use over the last 30 days.

For whites, the remarkable feature of these data is the extent of similarity between the two modes. The net differences between modes were zero or near zero in all six measures of marijuana use. For the state of New Jersey, then, there would be no change in marijuana use estimates for whites associated with a shift from face-to-face to telephone interviewing. The prediction of greater social desirability effects for the telephone versus personal interviewing was supported only by the data for black, but not for white, marijuana use.

Cocaine. The direction of mode effects for cocaine was similar to that found for alcohol and marijuana, although mode effects were not significant. The telephone survey provided lower estimates than the face-to-face survey of blacks" lifetime, last year, and current cocaine use. For whites, lifetime use estimates were higher by telephone, while last year and current use estimates were lower by telephone. The lone significant main effect was for whites: the number admitting cocaine use more than 10 times in their life was higher in the telephone survey (17% to 10%). Overall, then, the telephone mode produced generally, but not significantly, lower cocaine use estimates for blacks, and similar or higher estimates for whites.

Racial Differences by Interview Mode. Given the mode effects described above, it is important to ask whether interview mode might affect findings concerning racial differences in drug use. Net differences between races (black minus white estimates) are presented by interview mode in Table 4.

Although no significant mode effects for smoking were found, the smoking estimates for blacks were slightly higher by telephone, while for whites the smoking estimates were slightly lower. The effect of these trends can be seen in the net racial differences for tobacco. In the personal survey, smoking estimates were significantly higher for

Table 3. Weighted Drug Use Estimates, by Respondent Race and Interview Mode

	Black			White		
	Telephone	Personal	Net Difference[a]	Telephone	Personal	Net Difference[a]
	Tobacco (cigarette smoking)					
Smoked 5 packs or more lifetime (%)	42	43	−2	51	49	+2
Smoked in last 12 months (%)	42	37	+4	41	42	−2
Smoked in last 30 days (%)	39	33	+4	35	37	−3
No. of cigarettes smoked per day	4.4	3.6	+.3	5.5	6.2	−.7
No. of years smoked daily	3.5	3.2	−.1	3.7	3.6	0
	Alcohol					
Drank in last 30 days (%)	52	60	−8	77	84	−6**
No. of drinking days/last 30 days	1.8	4.0	−2.3**	5.3	5.5	−.2
Total drinks/last 30 days	4	16	−12**	18	18	0
No. of days had 5 or more drinks	0.3	1.6	−1.4*	1.3	1.5	−.3
Drunk once or more last year (%)	26	28	−2	53	59	−6*

	Marijuana					
Used once or more lifetime (%)	51	64	−16*	66	66	0
Used last 12 months (%)	16	22	−6	25	26	0
Used last 30 days (%)	7	17	−12*	13	13	0
Used 10 or more times in life (%)	17	31	−16**	39	39	0
No. of days used/last 30 days	0.6	1.4	−1.1*	1.0	1.1	−.1
No. of joints a day/last 30 days	0.1	0.5	−.6*	0.3	0.2	+.1
	Cocaine					
Used once or more lifetime (%)	19	23	−4	33	30	+4
Used last 12 months (%)	9	11	−2	12	14	−2
Used last 30 days (%)	3	5	−2	5	6	−2
Used 10 or more times in life (%)	6	9	−4	17	11	+5**
No. of days used/last 30 days	0.4	0.2	+.1	0.3	0.2	+.1

[a] Net differences (telephone − personal) controlling for sex, age, employment, income, education, student status, and marital status. Nontelephone households excluded from the personal survey.

** $p < .01$, * $p < .05$.

Table 4. Race Differences in Weighted Drug Use Estimates, by Interview Mode

	Telephone			Personal		
	Black	White	Net Difference[a]	Black	White	Net Difference[a]
	Tobacco (cigarette smoking)					
Smoked 5 packs or more lifetime (%)	42	51	−13**	43	49	−10*
Smoked in last 12 months (%)	42	41	−4	37	42	−10*
Smoked in last 30 days (%)	39	35	+1	33	37	−9*
No. of cigarettes smoked per day	4.4	5.5	−1.7#	3.6	6.2	−2.3***
No. of years smoked daily	3.5	3.7	−.5	3.2	3.6	−.7
	Alcohol					
Drank in last 30 days (%)	52	77	−23***	60	84	−24***
No. of drinking days/last 30 days	1.8	5.3	−3.2***	4.0	5.5	−1.3*
Total drinks/last 30 days	4	18	−13***	16	18	−3
No. of days had 5 or more drinks	0.3	1.3	−.9**	1.6	1.5	+.2
Drunk once or more last year (%)	26	53	−27***	28	59	−31***

	Marijuana					
Used once or more lifetime (%)	51	66	−17***	64	66	−1
Used last 12 months (%)	16	25	−10*	22	26	−4
Used last 30 days (%)	7	13	−8*	17	13	+5
Used 10 or more times in life (%)	17	39	−24***	31	39	−5
No. of days used/last 30 days	0.6	1.1	−.5	1.4	1.1	+.2
No. of joints a day/last 30 days	0.1	0.3	−.3#	0.5	0.2	+.3*
	Cocaine					
Used once or more lifetime (%)	19	33	−15***	23	30	−7
Used last 12 months (%)	9	12	−5	11	14	−4
Used last 30 days (%)	3	5	−2	5	6	−1
Used 10 or more times in life (%)	6	17	−12**	9	11	−5
No. of days used/last 30 days	0.4	0.3	+.1	0.2	0.2	−.1

[a] Net differences (Black − White) controlling for sex, age, employment, income, education, student status, and marital status. Nontelephone households excluded from the personal survey.

***$p < .001$. **$p < .01$. *$p < .05$. # $p = .06$.

whites on four of the five indicators. The racial differences were attenuated in the telephone survey, with only two estimates suggesting higher levels of smoking for whites.

For alcohol, marijuana, and cocaine, the number and size of race differentials were larger in the telephone than in the personal survey, with black estimates lower than white estimates in nearly all cases. The telephone survey furnished large and significant race differences for all five measures of alcohol consumption, compared to three of five measures in the personal survey. The race findings are most striking for marijuana use. The personal survey indicated significant race differences for only one of the six indicators, with blacks smoking a slightly higher number of joints per day than whites. The telephone estimates show black marijuana use significantly lower than whites' on five of the six indicators, including the use of significantly fewer joints per day than whites. Finally, the personal survey indicated no significant race differentials in cocaine use; by telephone, however, blacks appeared significantly less likely than whites to have ever used cocaine, or to have used it 10 times or more in their lifetimes. It is clear that the two interview modes compared in this study provide very different pictures of black-white differences in all four drug use categories.

Race-of-Interviewer Effects. Survey literature concerning racially sensitive questions has consistently demonstrated race-of-interviewer effects (Schuman and Kalton, 1985). When interviewed by someone of a different race, respondents tend to answer racially sensitive questions in ways considered favorable to the interviewer (Campbell, 1981). Blacks interviewed by whites display less hostility toward whites, less militant civil rights attitudes (Schaeffer, 1980; Schuman and Converse, 1971), and less closeness to blacks (Anderson et al., 1988b), than those interviewed by blacks. Similarly, whites interviewed by blacks respond more liberally on racial issues than those interviewed by whites (Cotter et al., 1982; Hatchett and Schuman, 1975).

These studies suggest that lack of racial matching between interviewer and respondent increases the likelihood of socially desirable responding. Although drug use items do not concern race per se, they may be "racially sensitive." If respondents consider drug use to be socially undesirable, having an interviewer of another race may increase respondents' tendencies to conceal undesirable behavior. Thus, it was important to determine the extent to which the interview mode and race findings in this research were attributable to race-of-interviewer effects.

The number of respondents interviewed by black and white interviewers is shown in Table 5. In the personal survey, nearly all whites, and slightly more than half of blacks, were interviewed by whites.

Table 5. Race of Respondent and Interviewer, by Mode of Interview

	Mode of Interview			
	Telephone		Personal	
	Race of Respondent			
Race of Interviewer	Black	White	Black	White
Black	76	427	58	26
	(73%)	(67%)	(36%)	(4%)
White	28	206	89	686
	(27%)	(32%)	(56%)	(91%)
Missing	—	3	13	40
		(.5%)	(8%)	(5%)
Total	104	636	160	752
	(100%)	(100%)	(100%)	(100%)

By telephone, a clear majority of both blacks (73%) and whites (67%) were interviewed by black interviewers. Thus, black respondents were more likely to be interviewed by someone of the same race in the telephone survey. White respondents were much more likely to have a same race interviewer in the personal survey. If racial matching (Schaeffer, 1980) of interviewer and respondent is assumed to elicit more accurate responses to sensitive questions, the telephone mode should have favored blacks in this study and the personal mode should have favored whites. The results reported in Table 3, however, do not support this assumption. Telephone estimates of drug use for whites, the majority of whom were interviewed by blacks, were nearly identical to estimates from the personal survey, where nearly all whites were interviewed by whites. Further, the increased racial matching of respondent and interviewer for blacks by telephone did not result in higher reporting of drug use than in the personal survey, where proportionately fewer blacks were interviewed by black interviewers.

The ability of racial matching to reduce socially desirable responses is controversial. Anderson, Silver, and Abramson (1988a) found blacks *more likely* to overreport voting when interviewed by a black rather than white interviewer, that is, racial matching produced less valid results for blacks. To verify that the patterns of drug self-reporting described in this paper are independent of interviewer race, we calculated drug use estimates for respondents interviewed by same race interviewers only. These estimates are displayed in Table 6. The mode

Table 6. Drug Use Estimates by Interview Mode for Respondents with Same Race Interviewers

	Black Respondent/ Black Interviewer			White Respondent/ White Interviewer		
	Telephone	Personal	p^a	Telephone	Personal	p^a
N of cases	76	58		206	686	
	Tobacco (cigarette smoking)					
Smoked 5 packs or more lifetime (%)	43	52	n.s.	57	51	n.s.
Smoked in last 12 months (%)	39	56	.06	44	41	n.s.
Smoked in last 30 days (%)	36	52	.08	37	36	n.s.
No. of cigarettes smoked per day	3.3	4.9	n.s.	5.9	6.1	n.s.
No. of-years smoked daily	3.4	3.9	n.s.	4.9	3.9	.02
	Alcohol					
Drank in last 30 days (%)	51	69	.06	75	84	.01
No. of drinking days/last 30 days	1.6	3.2	.05	5.0	5.4	n.s.
Total drinks/last 30 days	4	14	.09	16	17	n.s.
No. of days had 5 or more drinks	0.2	1.0	.03	1.1	1.4	n.s.
Drunk once or more last year (%)	25	33	n.s.	50	57	.08

	Marijuana					
Used once or more lifetime (%)	46	65	.04	68	67	n.s
Used last 12 months (%)	17	33	.04	21	25	n.s.
Used last 30 days (%)	8	25	.01	10	13	n.s.
Used 10 or more times in life (%)	17	31	.07	38	38	n.s.
No. of days used/last 30 days	0.6	2.2	.05	0.8	1.1	n.s.
No. of joints a day/last 30 days	0.1	1.3	.03	0.4	0.2	n.s.
	Cocaine					
Used once or more lifetime (%)	21	21	n.s.	33	29	n.s.
Used last 12 months (%)	12	13	n.s.	11	13	n.s.
Used last 30 days (%)	5	6	n.s.	4	6	n.s.
Used 10 or more times in life (%)	8	2	n.s.	15	11	n.s.
No. of days used/last 30 days	0.6	0.2	n.s.	0.2	0.2	n.s.

---

[a] *F*-test used to determine significance of mode effects; n.s. = nonsignificant.

effects on self-reports of drug use for respondents with same-race interviewers were nearly identical to those found for the full sample. For blacks, there were significantly lower alcohol and marijuana estimates by telephone, and some evidence of lower tobacco use estimates as well. For whites, the sole discrepancy between the full sample and those interviewed by whites was for years of daily smoking (a significantly higher estimate by telephone). The lower telephone estimates of drug use for blacks and the increased racial differences in drug use in the telephone versus personal survey do not appear to be attributable to differences between modes in the proportion of black and white interviewers, or to the racial matching of respondents and interviewers.

We should note, however, that interviewers were not randomly assigned to respondents. In the personal survey, black interviewers were assigned to black areas wherever possible and white interviewers to white areas. Respondents interviewed by same-race interviewers have different characteristics than respondents interviewed by someone of a different race. The results in Table 6 cannot be generalized beyond the samples included in this study.

Discussion

Looking down the column of results in Table 3, a justifiable conclusion for whites is that the telephone survey yielded results comparable to face-to-face estimates across all drug groups. Where telephone estimates for whites were significantly lower—for the current use of alcohol and drunkenness—the percentage differences between modes were small. The RDD sample also reproduced the demographic profile for whites found in the personal mode. The dearth of mode effects for whites is especially noteworthy given the large sample sizes (over 1,300 cases) and therefore relatively high power to detect differences.

The conclusions for blacks are very different. The telephone survey introduced a significantly higher SES profile for blacks. After controlling for SES and other demographic characteristics, the telephone survey resulted in substantially lower estimates of blacks' current alcohol consumption and marijuana use, compared to the face-to-face results. The telephone interviews also suggested more and larger racial differences in drug use than the personal survey (Table 4). By telephone, black alcohol and marijuana use was significantly lower than whites' in nearly all categories. The personal survey indicated almost no racial differences in marijuana use, and fewer racial differences in the amount of recent alcohol consumption. The personal survey indicated no significant racial differences in cocaine use; the telephone estimates suggested that blacks were significantly less likely than whites to have ever used cocaine, or to have used it more than ten times. We recognize that neither provides a fully accurate description of black drug use or demographic characteristics. Minority men were substantially underrepresented in the surveys, accounting for only one-third of the respondents in both modes.

MODE EFFECTS BY DRUG CATEGORY

Differences between telephone and personal survey estimates were not found for tobacco use, the least threatening of the four drug groups. This is consistent with the mode comparison literature on health surveys and other less threatening topics (Herzog et al., 1983; Herzog and Rodgers, 1988; Groves, Miller, and Cannell, 1987). Mode effects (for blacks) were more evident as the sensitivity of the questions increased, with the strongest differences for recent and heavy drinking and marijuana use. Contrary to expectation, however, the findings were not as clear for cocaine as for the other three drug groups.

Although the telephone survey furnished consistently lower cocaine use estimates for blacks than did the personal survey, the mode differences were not significant. This was somewhat surprising since questions on cocaine should be at least as sensitive as questions on marijuana use. Mensch and Kandel (1988), for example, found significant underreporting of both marijuana and cocaine in their study of social desirability effects on NLSY drug estimates. The mode differentials for black cocaine use in this study may have reached significance if the sample sizes were larger (as in the NLSY). Studies with larger samples of respondents and with additional drug categories may be needed to sort out the differences in interview mode effects among various drugs.

ORIGIN OF THE INTERVIEW MODE AND RACE EFFECTS

Why would blacks display more sensitivity to mode of interview effects than whites? It is not possible to answer this question definitively with data from this study, or through recourse to the survey literature. We do know, however, that minorities in general appear more prone than whites toward socially desirable responding in sensitive surveys, regardless of interview characteristics. Mensch and Kandel (1988) reported that blacks and Hispanics were twice as likely as whites in the NLSY to underreport drug use and that, in national drug use surveys, estimates of blacks' illicit drug use are consistently lower than estimates for whites. Blacks also appear less likely than whites to report their own criminal behaviors in surveys (Hindelang et al., 1981; Huizinga and Elliott, 1984). It is possible that minorities are more suspicious than whites of surveys of illicit behavior, and may have less confidence than whites in guarantees of confidentiality. In turn, heightened suspicions may foster increased reactivity to variations in survey methodology among blacks (or other minorities), especially in surveys of illicit or socially sanctioned behavior.

Although this study cannot pinpoint the exact causes of the interview mode effects, a few tenable conclusions can be drawn. The notion that sample characteristics alone account for the mode effects is not supported by the data. We have focused on two sources of bias: RDD's exclusion of nontelephone households and differences in demographic profiles due to nonresponse. Exclusion of nontelephone households from the face-to-face sample and controls for respondent demographic characteristics did not reduce the interview mode and respondent race effects for alcohol, marijuana, or cocaine use. Although respondent characteristics differ in RDD and personal modes, they do not fully account for mode differences in self-reported drug use. Interviewer effects also appeared unrelated to the mode differences. Controlling for interviewer race did not alter the pattern of mode effects for either blacks or whites.

Apart from sampling issues and nonresponse, we suspect that characteristics of the interview situation itself play a role in altering responses to sensitive or threatening survey items. Mode differences in provision of privacy and anonymity are central issues in surveys concerning illegal or socially sanctioned behavior. The self-administered answer sheets for drug use reporting in the personal mode, where interviewers read the questions but do not see the answers, may have provided a greater degree of anonymity to respondents than the telephone interview, where the unseen interviewer hears the report directly. Although randomized response techniques (RRTs) have been developed for telephone surveys (Orwin and Boruch, 1982), it would be very difficult to use these to insure greater privacy in drug surveys. In order to branch to the appropriate follow-up question, the telephone interviewer must know the answer to the filter question; e.g., follow-up questions concerning recent marijuana use can be asked only if lifetime use of marijuana is admitted, and recorded by the interviewer.

Another aspect of questionnaire design that may be involved in the mode effects is redundancy. In the face-to-face interview, respondents recorded an answer to each question on the self-administered answer sheet, regardless of their history of using that particular drug. Persons who indicated that they never used marijuana, for example, were not branched out of the marijuana sequence at that point, but were asked to respond to each marijuana use question. In addition to standardizing the sequence of questions for all respondents, this technique gave respondents more time to reconsider their answers or to admit using a drug after having initially denied it. The high degree of redundancy involved in this method is impossible in a telephone interview. Pretesting made it abundantly clear that, to avoid an extraordinarily high break-off rate, telephone respondents needed to be branched out of inappropriate question sequences.

Changes in question format for telephone administration (two-step and bracketing procedures) are most likely not related to the mode and race differences. Significant effects were obtained for revised items as well as for items with identical formats in both survey modes. Items with identical formats that yielded significant mode and race effects included: number of drinking days, total drinks, and number of days with five or more drinks; lifetime marijuana and cocaine use, number of days of marijuana use, and number of marijuana joints smoked.

APPLICABILITY TO OTHER SURVEYS

In surveys concerned with sensitive or embarrassing topics, such as sexual behavior, sexually transmitted diseases, drugs, or criminal behavior, the sample characteristics associated with RDD versus area probability sampling may be highly correlated with the behavior the survey is trying to estimate. It is imperative, in sensitive surveys, to correct for the nontelephone household exclusion and for the higher SES profile produced by RDD sampling. In addition, our findings suggest that the impact of social desirability on responses to sensitive questions may vary by mode of interview, and that mode of interview effects may themselves vary among racial and ethnic groups. The results of surveys whose purpose is to estimate racial and ethnic group differences in sensitive or illegal behavior may vary substantially by interview mode. If underreporting of illicit or embarrassing behavior is the biggest threat to validity, researchers interested in sensitive and threatening topics should not assume that all survey modes will furnish equivalent estimates.

Appendix: Verbatim Drug Use Questions

CIGARETTES—SAME WORDING IN TELEPHONE AND FACE-TO-FACE SURVEY

C1. The first question is about smoking tobacco. About how old were you when you first tried a cigarette? (Record age, or never tried a cigarette.)

C2. Have you smoked as many as five packs of cigarettes, that is, at least 100 cigarettes, in your life?

C3. When was the most recent time you had a cigarette? (In the past 30 days, within the past 6 months, within the past year, one to three years ago, more than three years ago, not sure.)

C4. How many cigarettes have you smoked, on the average, during the past 30 days? Give me the average number per day. (Less than one cigarette a day, one to five cigarettes a day, about half a pack a day [6–15 cigarettes], about a pack a day [16–25 cigarettes], about 1½ packs a day [26–35 cigarettes], about two packs or more a day [over 35 cigarettes], not sure.)

C5. For how many years did you smoke daily?

ALCOHOL—FACE-TO-FACE SURVEY

A1. The next questions are about alcohol, that is, beer, wine, and liquor, like whiskey, gin, or mixed alcoholic drinks like gin and tonic. About how old were you the first time you had a glass of beer or wine or a drink of liquor, such as whiskey, gin, scotch, or something else? Do not include childhood sips you may have had from an older person's drink. If you can't remember exactly how old you were, give your best guess. (Record age, or never had a drink.)

A2. When was the most recent time that you had an alcohol drink, that is, of beer, wine, or liquor or mixed alcoholic drinks? (Within the past month [30 days], within the past six months but more than a month ago, six months to a year ago, more than a year to three years ago, more than three years ago, never had a drink.)

A3. If you used alcohol during the past 30 days, on how many different days did you have one or more drinks?

A4. On the days that you drank during the past 30 days, about how many drinks did you usually have a day? By a drink we mean the equivalent of a can of beer, a glass of wine, a shot glass of hard liquor, or a mixed drink.

A5. During the past 30 days, about how many days did you have five or more drinks on the same occasion? By occasion we mean at the same time or within a couple of hours of each other.

ALCOHOL—TELEPHONE SURVEY

A1. The next questions are about alcohol, that is, beer, wine, and liquor, like whiskey, gin, or mixed alcoholic drinks like gin and tonic. About how old were you the first time you had a glass of beer or wine or a drink of liquor, such as whiskey, gin, scotch, or something else? Do not include childhood sips you may have had from an older person's drink. If you can't remember exactly how old you were, give your best guess. (Record age, or never had a drink.)

A2a. Was *the most recent time* you had an alcohol drink, that is, of beer, wine, liquor, or mixed alcoholic drinks: within the past 30 days, within the past 6 months but more than 30 days ago, or more than 6 months ago? (If more than six months ago, go to A2b.)

A2b. Was the most recent time you had an alcoholic drink: six months to a year ago, more than a year to 3 years ago, or more than three years ago?

A3. During the past 30 days, on how many different days did you have one or more drinks?

A4. On the days that you drank during the past 30 days, about how many drinks did you usually have a day? By a drink we mean the equivalent of a can of beer, a glass of wine, a shot glass of hard liquor, or a mixed drink.

A5. During the past 30 days, about how many days did you have five or more drinks on the same occasion? By occasion we mean at the same time or within a couple of hours of each other.

A6a. In the past 12 months, have you gotten very high or drunk on alcohol more than twice a month, that is, at least 25 days a year?

A6b. (If yes to A6a): In the last 12 months, have you gotten very high or drunk on alcohol: daily, 3 to 6 days a week, about 1 or 2 days a week, or several times a month, but less than once a week?

A6c. (If no to A6a): In the last 12 months, have you gotten very high or drunk on alcohol: 1 or 2 times a month, every other month or so, or 6 to 11 days a year, 3 to 5 days in the past 12 months, 1 or 2 days in the past 12 months, or not at all?

MARIJUANA—FACE-TO-FACE

M1. About how old were you the first time you actually used marijuana or hash? (Record age, or never.)

M2. Think of all the times you've actually used marijuana, from the first time up to the most recent time. About how many times in your life have you used marijuana or hash? (1 or 2 times, 3 to 5 times, 6 to 10 times, 11 to 49 times, 50 to 99 times, 100 to 199 times, 200 or more times, never.)

M3. When was the most recent time you used marijuana or hash? (Within the past week, within the past month [30 days] but more than a week ago, within the past six months but more than a month ago, six months to a year ago, more than a year to three years ago, more than three years ago, never.)

M4. The next few questions are about the last 30 days. On about how many different days did you use marijuana or hash during the past 30 days?

M5. The rest of these questions are only about marijuana and not hash. About how many marijuana cigarettes, joints, or reefers did you smoke per day on the average during the past 30 days? If you shared them with other people, count only the amount you smoked.

MARIJUANA—TELEPHONE

M1. About how old were you the first time you actually used marijuana or hash? (Record age, or never.)

M2a. Next, think about all the times you've used marijuana, from the first time up to the most recent time. Have you used marijuana or hash 50 times or more in your life? (If no, skip to M2d.)

M2b. In your life, have you used marijuana or hash 200 times or more? (If yes, skip to M3a.)

M2c. In your life, have you used marijuana or hash 100 times or more? (Skip to M3a.)

M2d. In your life, have you used marijuana or hashish: 1 or 2 times, 3 to 5 times, 6 to 10 times, or more than 10 times?

M3a. Think about the *most recent time* you used marijuana or hashish. Was the most recent time within the last six months?

M3b. (If yes to M3a): Was the most recent time you used marijuana or hashish: within the past week, within the past 30 days, but more than a week ago, or was it more than 30 days ago?

M3b. (If no to M3a): Was the most recent time you used marijuana or hashish: six months to a year ago, more than a year to three years ago, or more than three years ago?

M4. The next few questions are about the last 30 days. On about how many different days did you use marijuana or hash during the past 30 days?

M5. The rest of these questions are only about marijuana and not hash. About how many marijuana cigarettes, joints, or reefers did you smoke per day on the average during the past 30 days? If you shared them with other people, count only the amount you smoked.

COCAINE—FACE-TO-FACE

CN1. About how old were you the first time you actually used cocaine? (Record age, or never.)

CN2. Think of all the times you've used cocaine, from the first time up to the most recent time. About how many times in your life have you used cocaine? (1 or 2 times, 3 to 5 times, 6 to 10 times, 11 to 49 times, 50 to 99 times, 100 to 199 times, 200 or more times, never.)

CN3. When was the most recent time you used cocaine? (Within the past week, within the past month [30 days] but more than a week ago, within the past six months but more than a month ago, six months to a year ago, more than a year to three years ago, more than three years ago, never.)

CN4. The next few questions are about the past 30 days. During the past 30 days, on about how many different days did you use cocaine?

COCAINE—TELEPHONE

CN1. About how old were you the first time you actually used cocaine? (Record age, or never.)

CN2a. Next, think about all the times you've used cocaine, from the first time up to the most recent time. Have you used cocaine 50 times or more in your life? (If no, skip to CN2d.)

CN2b. In your life, have you used cocaine 200 times or more? (If yes, skip to CN3a.)

CN2c. In your life, have you used cocaine 100 times or more? (If yes, skip to CN3a.)

CN2d. In your life, have you used cocaine: 1 or 2 times, 3 to 5 times, 6 to 10 times, or more than 10 times?

CN3a. Was the most recent time you used cocaine within the past 12 months?

CN3b. (If yes to CN3a): Was the most recent time you used cocaine: within the past 30 days, within the past six months, but more than 30 days ago, or was it more than six months ago?

CN3c. (If no to CN3a): Was the most recent time you used cocaine: six months to a year ago, more than a year to three years ago, or more than three years ago?

CN4. The next few questions are about the past 30 days. During the past 30 days, on about how many different days did you use cocaine?

References

Anderson, B. A., B. D. Silver, and P. R. Abramson (1988a)
"The effects of race of the interviewer on measures of electoral participation by blacks in SRC national election studies." Public Opinion Quarterly 52:53–83.

——— (1988b)
"The effect of race of the interviewer on race-related attitudes of black respondents in SRC/CPS national election studies." Public Opinion Quarterly 52:289–324.

Bradburn, N. M., S. Sudman, E. Blair, and C. Stocking (1978)
"Question threat and response bias." Public Opinion Quarterly 42:221–234.

Campbell, B. A. (1981)
"Race-of-interviewer effects among southern adolescents." Public Opinion Quarterly 45:231–244.

Colombotos, J. (1969)
"Personal versus telephone interviews: Effect on responses." Public Health Reports 84:773–782.

Cotter, P. R., J. Cohen, and P. B. Coulter (1982)
"Race-of-interviewer effects in telephone interviews." Public Opinion Quarterly 46:278–284.

Freeman, H., K. Kiecolt, W. Nicholls, and J. M. Shanks (1982)
"Telephone sampling bias in surveying disability." Public Opinion Quarterly 46:392–407.

Groves, R. M. (1979)
"Actors and questions in telephone and personal interview surveys." Public Opinion Quarterly 43:190–205.

——— (1983)
"Implications of CATI: Costs, errors, and organization of telephone survey research." Sociological Methods and Research 12:199–215.

Groves, R. M., and R. L. Kahn (1979)
Surveys by Telephone: A National Comparison with Personal Interviews. New York: Academic Press.

Groves, R. M., P. V. Miller, and C. F. Cannell (1987)
"Differences between the telephone and personal interview data." In O. Thornberry (ed.), An Experimental Comparison of Telephone and Health

Interview Surveys. Vital and Health Statistics, series 2, no. 106. DHHS Pub. No. (PHS)87-1380. Washington, DC: Public Health Service.

Hatchett, S., and H. Schuman (1975)
"White respondents and race-of-interviewer effects." Public Opinion Quarterly 39:523–528.

Hauck, M., and M. Cox (1974)
"Locating a sample by random digit dialing." Public Opinion Quarterly 38:253–260.

Henson, R., C. F. Cannell, and A. Roth (1978)
"Effects of interview mode on reporting of moods, symptoms, and need for social approval." Journal of Social Psychology 105:123–129.

Herman, J. B. (1977)
"Mixed-mode data collection: Telephone and personal interviewing." Journal of Applied Psychology 62:399–404.

Herzog, A. R., and W. L. Rodgers (1988)
"Interviewing older adults: Mode comparison using data from a face-to-face survey and a telephone resurvey." Public Opinion Quarterly 52:84–99.

Herzog, A. R., W. L. Rodgers, and R. A. Kulka (1983)
"Interviewing older adults: A comparison of telephone and face-to-face modalities." Public Opinion Quarterly 47:405–418.

Hindelang, M. J., T. Hirschi, and J. G. Weis (1981)
Measuring Delinquency. Beverly Hills, CA: Sage.

Hochstim, J. R. (1967)
"A critical comparison of three strategies of collecting data from households." Journal of the American Statistical Association 62:976–989.

Huizinga, D., and D. S. Elliott (1984)
Self-Reported Measures of Delinquency and Crime: Methodological Issues and Comparative Findings. Boulder, CO: Behavioral Research Institute.

Johnston, L., and P. O'Malley (1985)
"Issues of validity and population coverage in student surveys of drug use." In B. Rouse, N. Kozel, and L. Richards (eds.), Self Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity. NIDA Research Monograph #57. Washington, DC: National Institute on Drug Abuse.

Jordan, L., A. Marcus, and L. Reeder (1980)
"Response styles in telephone and household interviewing: A field experiment." Public Opinion Quarterly 44:210–222.

Kish, L. (1965)
Survey Sampling. New York: John Wiley.

Klecka, W. R., and A. J. Tuchfarber (1978)
"Random digit dialing: A comparison to personal surveys." Public Opinion Quarterly 42:105–114.

Locander, W. B., and J. P. Burton (1976)
"The effect of question form on gathering income data by telephone." Journal of Marketing Research 13:189–192.

Mensch, B. S., and D. B. Kandel (1988)
"Underreporting substance use in a national longitudinal youth cohort: Individual and interviewer effects." Public Opinion Quarterly 52:100–124.

Miller, P. (1984)
"Alternative question forms for attitude scale questions in telephone interviews." Public Opinion Quarterly 48:766–778.

O'Neil, M. (1979)
"Estimating the nonresponse bias due to refusals in telephone surveys." Public Opinion Quarterly 43:218–232.

Orwin, R. G., and R. F. Boruch (1982)
"RRT meets RDD: Statistical strategies for assuring response privacy in telephone surveys." Public Opinion Quarterly 46:560–571.

Rogers, T. F. (1976)
"Interviews by telephone and in person: Quality of responses and field performance." Public Opinion Quarterly 40:51–65.

Schaeffer, N. C. (1980)
"Evaluating race-of-interviewer effects in a national survey." Sociological Methods and Research 8:400–419.
Schmiedeskamp, J. (1962)
"Reinterviews by telephone." Journal of Marketing 26:28–34.
Schuman, H., and J. M. Converse (1971)
"The effects of black and white interviewers on black responses in 1968." Public Opinion Quarterly 35:44–68.
Schuman, H., and G. Kalton (1985)
"Survey methods." In G. Lindzey and E. Aronson (eds.), The Handbook of Social Psychology, Vol. 1: Theory and Method. New York: Random House.
Siemiatycki, J. (1979)
"A comparison of mail, telephone, and home interview strategies for household health surveys." American Journal of Public Health 69:238–244.
Sudman, S. (1976)
"Sample surveys." Annual Review of Sociology 2:107–120.
Sudman, S., and N. Bradburn (1974)
Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine Publishing.
Thornberry, O. T., and J. T. Massey (1988)
"Trends in United States telephone coverage across time and subgroups." In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nichols, and J. Waksberg (eds.), Telephone Survey Methodology. New York: John Wiley.
Tull, D. S., and G. Albaum (1977)
"Bias in random digit dialed surveys." Public Opinion Quarterly 41:389–395.
Waksberg, J. (1978)
"Sampling methods for random digit dialing." Journal of the American Statistical Association 73:40–46.
Weaver, C., S. Holmes, and N. Glenn (1975)
"Some characteristics of inaccessible respondents in a telephone survey." Journal of Applied Psychology 60:260–262.
Wiseman, F. (1972)
"Methodological bias in public opinion surveys." Public Opinion Quarterly 36:105–108.

THE OBSESSION TO FORECAST: PRE-ELECTION POLLS IN THE ISRAELI PRESS

GABRIEL WEIMANN

Abstract The generous coverage of pre-election polls in the media has stimulated a debate on the effects of publicized polls on voters. Data for this study come from a content analysis of all press reports on polls during six pre-election campaigns in Israel (1969–88). The findings highlight the impressive increase of poll reports both by space and by prominence in the press. Analysis of the reports' content reveals that the style of reporting polls has become dominated by "horse-race" journalism, where predictions and popularity ratings are the leading themes. In addition, although the reports have become more detailed and informative, they still do not adequately define or interpret methodological deficiencies. Finally, the analysis examines the intercorrelations between attributes of coverage and prediction errors. The fact that the independent variable, media coverage of polls, has changed significantly calls for the inclusion of this factor in any study of polls' effects.

The generous coverage of pre-election polls in the media has aroused interest in the effects of such media coverage on voters' preferences. A recent survey of studies on these effects (Donsbach, 1987) reveals the lack of consistency in empirical evidence regarding the nature and the magnitude of the polls' influence. Several studies and reviewers of public opinion research argue that there is no significant change in public attitudes or voting behavior (e.g., Atkin, 1969; Klapper, 1960; Mendelsohn and Crespi, 1970; Roshwalb and Resnicoff, 1971) while the "strong effect" claimers find support in the ideas and findings of Noelle-Neumann (1973, 1974, 1976, 1984), calling for a "return to the concept of powerful mass media."

The reason for this lingering debate and the confusing findings may

GABRIEL WEIMANN is Associate Professor of Communications and Chair of the Department of Sociology at Haifa University, Israel, and Visiting Professor at the Institut für Publizistik, Universität Mainz, Germany. The study was supported by a research grant from the Alexander von Humboldt-Stiftung, West Germany.

Public Opinion Quarterly Volume 54:396–408 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

be a combination of several factors such as: (*a*) the general debate on media effects (see reviews by Katz, 1980 and 1988); (*b*) the multidimensionality of the effects (the political effects of publicized polls may involve cognitive, affective, or behavioral influences or various combinations of these dimensions); (*c*) short-term versus long-term effects; (*d*) direct and indirect effects; and (*e*) changes in media coverage of polls. The coverage of polls varies across societies and cultures, across political systems and regimes, and over time. This makes cross-cultural and longitudinal comparisons less valid. The polling industry has changed and so have the relationships between the media and the polling agencies (Gollin, 1988; Shamir, 1986). The fact that the indepedent variable itself, that is, the presentation of polls in the media, may have changed over time is another potential explanation for the inconsistency among studies that focus only on the effects. The present study focuses on the last factor: it examines the changes in press coverage of polls during six political campaigns in Israel.

Data

Data for this study come from a systematic content analysis of 15 daily newspapers in Israel during the pre-election campaigns of 1969, 1973, 1977, 1981, 1984, and 1988. The coders were trained to analyze any report on polls published during the three months preceding each election, using a preset manual of coding categories (see Charlebois, 1979; Stone and Morrison, 1976). The following components were recorded for each report:

a) Form of coverage: space allocation, placement in the newspaper, and source of sponsorship. Space was measured by column inches devoted to reporting a poll and its results and to analysis or interpretation of polls.

b) Content of coverage: inclusion of information on sampling method and sample size, wording of questions, rate of "don't know" and "undecided" answers, method of interviewing, timing, measures of intensity of opinion, and identification of pollster.

c) Style of coverage: "horse-race" polls (focusing on predictions), personal popularity polls, and issue-related polls (focusing on issues and attitudes).[1]

---

[1] The coders classified the statements in each article. Statements referring to distribution of votes, parliament seats, or party support were classified as "prediction statements." References to public opinion regarding issues (e.g., the occupied territories, talks with the PLO, inflation, security, and others) were coded as "issues/attitudes

The coders were trained prior to the investigation and were subjected to a reliability test. This test involved the comparison of the evaluations of different coders when analyzing the same text (see Holsti, 1969:135–142). The coefficients of reliability (i.e., the average proportion of agreement between pairs of coders) were .94 in form, .88 in content, and .89 in style. Parts of this data base were used for previous studies (Weimann, 1983a and 1983b) but for the purpose of the present study new measures were added, as well as the 1984 and 1988 campaign data.

Findings

DIFFUSION OF POLL REPORTS

The Israeli press began reporting pre-election polls only during the election campaign for the seventh Knesset in 1969. None of those polls were initiated or commissioned by the press; most reports originated with political bodies that leaked partial results to the press. However, subsequent campaigns witnessed a dramatic change in press coverage of polls. Table 1 presents several measures that demonstrate this change.

By any measure, polls have become an important news item, gaining more and more attention, space, and prominence in the press. The data reveal the gradual increase in the number of dailies covering polls (from 3 in 1969 to 15 in 1988[2]), in the number of reports on polls during the three months preceding the elections (from 16 in 1969 to 409 in 1988 when, on the average day, 4.5 polls were published every day), in space allocation (from 172 column inches in 1969 to 9365 in 1988), and in prominent placement (from 7% on the front page in 1969 to 25% in 1988, percentages are from all reports on polls). The growing popularity of polls in the Israeli press is also revealed in the frequency of reporting: the average number of reports per newspaper increased from 5.3 reports per paper in 1969 to 19.6 in 1981 and 27.2 in 1988. However, the dailies varied in the frequency of reporting, with the

---

statements." References to popularity of persons in terms of public support (usually references to the party leaders or candidates for the positions of Prime Minister and ministers) were coded as "personal statements." Statements combining more than one theme were classified in two (or more) categories. Then, each article was classified by its dominant theme, that is, the most frequent type of statement. Analysis was carried out both on the statement level as well as on the article level.

[2] This includes all eight Israeli dailies published in Hebrew, one in English (the *Jerusalem Post*), two in Arabic, and five in European languages such as Hungarian, Russian, Rumanian, French, and Yiddish. While the first papers to report polls were the Hebrew dailies, the rest joined later, mainly during the 1977 and 1981 campaigns.

Table 1. Measures of Polls Coverage, 1969–1988

	1969	1973	1977	1981	1984	1988	Rate of Change 69–73 to 84–88
Number of dailies reporting polls	3	6	8	11	13	15	3.11
Number of reports on polls	16	52	109	216	308	409	10.54
Average number of reports per newspaper	5.3	8.7	13.6	19.6	23.7	27.2	3.63
Space allocated for polls (col. in.)	172	513	1297	6927	8122	9365	25.22
Placement							
Front page (%)	7	12	15	21	24	25	3.47
Inside page (%)	93	88	85	79	76	75	−1.19
Sponsorship							
Reporting paper (%)	0	9	37	47	48	48	10.66
Other papers (%)	5	22	28	33	32	30	2.99
Political bodies (%)	71	49	22	11	12	16	−4.28
Unspecified (%)	24	20	13	9	8	6	−3.14

major dailies publishing the results of polls nearly every other day (in 1984–1988); at times reports on several (and even contradictory) polls were published in a single day's issue. During this period the Israeli press remained relatively stable in terms of the number of dailies, total space, and formats.

The growing trend to report polls has caused changes in the relationships between the press and polling agencies. *Ha'aretz* was the first paper to sponsor private polls (in 1973). Soon, the three major dailies commissioned their own polls, each contracting a different pollster. Consequently, while in 1973 only 9% of the reports covered polls sponsored by the reporting paper, the rate increased to 37% in 1977 and to 48% in 1984 and 1988. Moreover, the coverage of press-initiated polls was not confined to self-sponsored projects; all the newspapers included reports on polls sponsored by other papers. This trend yielded a growing share of press-commissioned polls which reached a peak in the 1980s: about 80% of the publicized polls were commissioned either by the reporting paper or by other papers.

The decline in reporting polls sponsored by political parties may be explained by the growing reluctance of politicians to leak such results due to "strategic" considerations[3] or by the growing scepticism of journalists to rely on this type of data.[4]

METHODOLOGICAL DEFICIENCIES

The emergence and prevalence of polls in the Israeli press is manifest not only in the changing form and prominence of poll reports but also in the contents of the reports. The comparisons presented in Table 2 indicate that content elements have changed significantly.

In the 1970s press reports of pre-election polls in Israel read merely as a short summary of results and predictions. But during the subsequent campaigns the reports have become more richly textured and detailed. Nevertheless, as Table 2 reveals, such enrichment has been rather selective. Comparison over the years shows that the reports have become more informative about sample size, definition of population, and identification of pollster and sponsor. To a lesser extent, an increase has also been noted in reports containing information regarding the method of interviewing, wording of questions, and timing. However, the reports still do not adequately define or interpret meth-

---

[3] Typical of this reluctance was the case of a pre-election poll conducted during May, 1977, for the Likud Party. Likud leaders decided to conceal the poll, whose results forecast their surprising victory. Its existence was publicized only after the elections.

[4] This scepticism was encouraged by several cases of false or manipulated polls leaked to the press. The worst case was that of a poll published by *Yediot* (February 15, 1974), later discovered to be a hoax, deliberately "planted" by a political party.

Table 2. Contents of Polls Reports, in Percentage of Reports Including Item

Item	1973	1977	1981	1984	1988	U.S. Papers 1972–79[a]
Sample size	54	86	92	90	90	89
Definition of population	56	68	84	80	76	91
Sampling error	0	6	9	8	8	31
Method of interviewing	29	32	51	50	52	62
Wording of question	14	21	28	35	36	71
Refusal rate	63	71	88	80	67	—
Timing of poll	38	51	64	66	65	76
Intensity of opinion	0	8	11	10	10	—
Identification of pollster	75	87	96	95	94	—
Identification of sponsor	30	49	66	76	88	—

[a] Data for this comparison come from a study of poll reports in the *Chicago Tribune*, the *Los Angeles Times*, and the *Atlanta Constitution*, conducted by Miller and Hurd (1982).

odological intricacies such as sampling error, rates of "don't know" or "undecided" answers, or measures of the intensity of opinion. The case with sampling problems is more crucial in the Israeli context since most of the pollsters use samples drawn from the urban Jewish population, thus misrepresenting nonurban and non-Jewish voters (e.g., the Arab voters in Israel are rarely represented in these polls). The rates of the undecided, "don't know," and "no answer" in the Israeli polls comprise 25% to 30%[5] of the surveyed samples, thus further weakening the validity of the predictions. Yet, this information is not included in a third of the reports on polls.

The data in Table 2 show that press reports on polls have become more detailed and informative, but not about the methodological limitations that may alter the conclusions derived from the results. This tendency was noted by Bogart, who claimed that "public misunderstanding of opinion surveys can be expected to continue as long as the mass media ignore or belittle their technical intricacies" (Bogart, 1972:23) and by Wheeler, who pointed to the manipulation of polls, arguing that "the reader of opinion polls is given no hint that they may not be trustworthy" (Wheeler, 1976:16). A study of poll reports in the

---

[5] This high rate was almost constant in all the campaigns of the 1970s and 1980s. It can be related to the fact that there are so many parties involved (26 in 1988) and that many of the undecided are actually "floating" between the big parties and their satellite parties. It is not related to the wording of the question, as all the Israeli pollsters apply the standard question (namely, "if the elections were held today . . .").

U.S. press led the researchers to the conclusion: "The way methodological information about polling is reported in the media tends more to reassure than alert the audience about the possible defects of poll data" (Paletz et al., 1980:506). It appears that these problems are more evident in the Israeli case. Comparison of the Israeli data with the findings of Paletz and his collaborators or with those of Miller and Hurd (1982) reveals that the Israeli press provides less adequate information about methodological deficiencies than does the American press. While 71% of the American reports include the polling question, the respective rate in the Israeli press is 36% (in 1988). Moreover, while 31% of the former include information on sampling errors, only 8% of the latter contain such information. Similar differences are found regarding the definition of the population (91% and 76% respectively) and method of interviewing (62% and 52% respectively). The fact that the Israeli polling industry has no agreed standards of reporting polls (such as the American Association for Public Opinion Research standards, see Gollin, 1988) seems to explain these differences. Thus, the combination of the growing prominence and sophistication of polls coverage in Israel and the accompanying stagnation in reporting the methodological intricacies may result in changing the impact these reports have on their readers' evaluations and preferences.

"HORSE-RACE JOURNALISM"

One of the popular criticisms of poll reports in the media is related to their focusing on superficial aspects of public opinion and politics. Rokeach noted this trend: "The kinds of data obtained by public opinion research and disseminated in the mass media seem designed more to entertain than to inform. The quality of the information conveyed seems not much different from that conveyed in the sports page, or better yet, the daily racing form" (Rokeach, 1968:548). Several studies (Broh, 1980; Paletz et al., 1980; Stovall and Solomon, 1983) revealed this tendency of poll coverage to deal exclusively with the "horse-race" aspects of the political campaign. Table 3 presents a comparison of types of polls reported in five campaigns. The reports were classified according to their main theme: predicting the elections' results (prediction polls), personal popularity polls, and polls that focus on issues and political attitudes.[6]

[6] In the Israeli campaigns the issues of security, the occupied territories, negotiation with the Palestinians, inflation, and economic stability were the leading themes and the major factors shaping the voters' preferences. Studies of the issues dominating the Israeli campaigns and their influence on voters (see for example Elizur and Katz, 1979; Katz, 1989) reveal the centrality of these issues in determining voting in Israel. Moreover, they indicate that the campaigns themselves focused on these issues with changing emphases from one campaign to another.

Table 3. Poll Reports According to Type

Type of Poll	1973	1977	1981	1984	1988	Rate of Change 73–77 to 84–88
Prediction polls	46	41	42	52	60	1.28
Personal popularity	8	22	24	28	28	1.86
Issues and attitudes	46	37	34	20	12	−2.59

Coverage of polls in the Israeli press has evidently displayed a consistent trend toward horse-race imagery. In every campaign, the "prediction polls" were the most popular polls reported and this preference has been increasing over the years (from 46% of the reports in 1973 to 60% in 1988). Moreover, there is a consistent increase in the "candidate-centered" reports, a trend more noteworthy in Israel where voters choose among parties and not candidates. The comparison in Table 3 reveals the decline of reports on issues and political attitudes (from 46% of the reports in 1973 to only 12% in 1988). Thus, we find that more and more polls reported in the press deal with predictions or with personalities and not with political issues. It appears that polls are used more often to rate parties and candidates than to study voters, the decision-making process, and the political issues involved. Broh argues that this type of predictions combined with personal ratings are the characteristics of the horse-race coverage of polls: "For journalists, the horse-race metaphor provides a framework for analysis. A horse is judged not by its absolute speed or skill but in comparison to the speed of other horses and especially by its wins and losses. Similarly, candidates are pushed to discuss other candidates; events are understood in a context of competiton; and picking the winner becomes an important topic. The race—not the winner—is important" (Broh, 1980:528).

ACCURACY AND ATTRIBUTES OF COVERAGE

How is the accuracy of polls to be measured? The fact that polls reflect the state of public opinion at the time they are conducted makes the measurement of accuracy almost impossible. One way is to examine interpoll consistency when comparing reports on polls taken at the same time: such comparisons in the Israeli context revealed significant interreport variance across reports on the same poll or polls taken at the same time (Shamir, 1986; Weimann, 1989). Another approach relies on defining accuracy as "the degree of congruence between pollsters'

Table 4. Correlations between Attributes of Coverage and Accuracy

	A	B	C	D	E
A. Prominence (space by column inches)	—	.07	−.54*	.36*	.05
B. Number of methodological deficiencies reported		—	.47*	−.21*	−.51*
C. Timing (number of days left until the elections)			—	−.62*	.46*
D. Percentage of horse-race statements (prediction + personal polls)				—	.08
E. Average prediction error					—

* $p < .01$ under the null hypothesis of no significant correlation.

final pre-election predictions and actual results" (Shamir, 1986:69). Although polls are presented as reflecting the temporary state of voting intentions they are often perceived by the media, the politicians, and the public as predicting the results of the elections.

The present paper applies the second approach to the measurement of accuracy by comparing each prediction published in a press report on polls, throughout the campaign, with the actual results. While Shamir (1986) studied only the final predictions, the present analysis examines all the published predictions, thus enabling the measurement of changes over time. Table 4 presents Spearman's rank-order correlations between measures of coverage and the overall error calculated by the mean deviation of polls' predictions—for the two leading parties and all the other parties together—from the actual results.[7]

The accuracy of reported polls, tested by the actual results of the elections, was found to be strongly related to two variables (see the last column of Table 4). It appears that the more methodological deficiencies are reported, the smaller is the prediction error ($R = -.51$). The reports that carefully detail the methodological problems, thus limiting the value of predictions based on the poll, are more accurate in their predictions. This reflects the professional standards of the pollster, which affect both the form of reporting his or her findings as well as the accuracy of his or her predictions. As expected, prediction errors decline with the number of days left to the elections ($R = .46$). One should also note the lack of correlation with prominence, indicating that polls gaining more coverage in the press are not necessarily more accurate in their predictions.

---

[7] For a detailed presentation of the computational formula as applied in the Israeli context, see Shamir 1986. It is based on collapsing the results of the poll into three categories (Likud, Alignment, and all others) and calculating the mean error, that is, the mean difference between the number of seats predicted for each of the three blocs and the actual number following the election outcomes. The published polls deviated by an average of about 8 seats in their predictions for the two parties with some polls deviating by more than 16 seats. This level of deviation is almost constant across the campaigns. The other variables were: prominence (space in column inches), timing (number of days between day of publication and election day), number of methodological deficiencies (references to problems like sampling, the undecided and DKs, statistical significance, underrepresentation of subpopulations, and others), and percent of horse-race statements. These measures were correlated with the prediction error and the resulting coefficients are presented in Table 4.

Finally, the intercorrelations between various attributes of coverage reveal interesting patterns. Timing is strongly correlated with all other variables: the closer the polls are to election day the more coverage is given ($R = -.54$) but this coverage becomes less detailed regarding the methodological problems ($R = .47$) and more devoted to the horse-race type of reporting ($R = -.62$). Thus, when the campaign is nearing its peak the "obsession to forecast" is more evident, making polls more newsworthy and prominent, but less cautious about polls' deficiencies and more focused on ratings and predictions than on political issues and attitudes. Prominence of poll reports is also found to be related to the horse-race genre: the reports that gain more coverage tend to be those that focus on personal or party ratings ($R = .36$). The horse-race reports tend to be less detailed regarding the methodological limits ($R = -.21$), thus providing the reader with more predictions but less information that limit the validity of these predictions.

Conclusions

Guided by the ongoing debate concerning the political impact of pre-election polls, this study examined the changes in the form and content of polls coverage during six election campaigns. The findings suggest both methodological and substantive implications. From the methodological perspective the necessity to include coverage analysis is highlighted. A host of inconsistencies among studies on the effects of published polls on the voters may be resolved by adding the variation in press coverage itself as a contributing factor.

From the substantive perspective, the findings highlight the growing tendency of the Israeli press to conduct and commission polls, involving the media more and more in the actual "newsmaking" process. This involvement may affect poll accuracy, either by a newspaper's endorsement that was found to determine the poll coverage (see Wilhoit and Sup Auh, 1974) or by exerting pressure on the pollsters. In his study on the accuracy of pre-election polls in the 1981 and 1984 Israeli campaigns, Shamir found that "Pollsters quest for more detail and refinement is driven by competition and mass media pressures for highly processed predictions is liable to result in less accurate reports" (Shamir, 1986:62).

Poll reports were found to change not only in amount and prominence but also in their contents. The consistent trend toward horse-race reporting, the selective enrichment of the reports, with methodological deficiencies omitted, the personification of political campaigns by turning them into "popularity contests" rated by polls—all may be harmful to the democratic process by affecting the voters ("Just as bettors on a race may be misled by the beauty of the horses—voters may be misled by the trivia of the campaign" Broh, 1980:528). Also affected are the politicians who are drawn to short-term decisions and actions that may result in better standing in the polls. The emergence of "precision journalism" (Meyer, 1973 and 1989) has led to growing acceptance of polls as both important and newsworthy (a fact that can be related to the accurate predictions of many American polls). However, when this growing reliance on surveys and polls is not accompanied by increasing familiarity and understanding of the statistical and methodological problems involved in polling, and when standards for reporting polls are nonexistent or poorly observed, the results of such "precision journalism" would emerge as far from accurate and valid.

The findings of the present study highlight the dramatic changes in the quantity, prominence, and contents of press reports on polls and in the relationships between the media and the polling industry in one society. The findings should be limited by the special circumstances of the Israeli campaigns, but nevertheless suggest that such monitoring of the changing presentation of polls in the media should be applied to other societies as well, testing Gollin's (1980:446) argument: "If the press—here meant to include both print and broadcast news media—has become more of a Fourth Estate in fact, then polls can be said to be one of its sharpest swords."

References

Atkin, C. A. (1969)
"The impact of political poll reports on candidates and issue preferences." Journalism Quarterly 46:515–21.

Bogart, L. (1972)
Silent Politics: Polls and the Awareness of Public Opinion. New York: John Wiley.

Broh, C. A. (1980)
"Horse-race journalism: Reporting the polls in the 1976 elections." Public Opinion Quarterly 44:514–29.

Crespi, I. (1980)
"Polls as journalism." Public Opinion Quarterly 44:462–76.
Charlebois, C. (1979)
"Multiple measures in the study of press response." Journalism Quarterly 56:851–56.
Donsbach, W. (1987)
"Are poll data hazardous to the health of democratic elections?" Paper presented at the annual conference of the International Communication Association (ICA), Montreal.
Elizur, J., and E. Katz (1979)
"The media in the elections of 1977." Pp. 227–254 in M. R. Penniman (ed.), Israel at the Polls. Washington DC: American Enterprise Institute for Public Policy Research.
Gollin, A. E. (1980)
"Exploring the liason between polling and the press." Public Opinion Quarterly 44:455–61.
——— (1987)
"Polling and the news media." Public Opinion Quarterly 51:86–94.
Holsti, O. R. (1969)
Content Analysis for the Social Sciences and Humanities. New York: Addison-Wesley.
Katz, E. (1980)
"On conceptualizing media effects." Studies in Communication 1:19–141.
——— (1987)
"Communication research since Lazarsfeld." Public Opinion Quarterly 51:25–45.
Katz, E., and P. Lazarsfeld (1955)
Personal Influence. Glencoe, IL: Free Press.
Katz, E., and H. Levinsohn (1989)
"Too good to be true: Notes on the Israeli elections of 1988." International Journal of Public Opinion Research 1 (2):111–22.
Klapper, J. (1960)
The Effects of Mass Communication. Glencoe, IL: Free Press.
Mendelsohn, H., and I. Crespi (1970)
Polls, Television and the New Politics. Scranton, PA: Chandler.
Meyer, P. (1973)
Precision Journalism. Bloomington, IN: Indiana University Press.
——— (1989)
"Precision journalism and the 1988 elections." Paper presented at the Esomar Conference on Public Opinion, Seville, Spain.
Miller, M. M., and R. Hurd (1982)
"Conformity to AAPOR standards in newspaper reporting of public opinion polls." Public Opinion Quarterly 46:243–49.
Noelle-Neumann, E. (1973)
"Return to the concept of powerful media." Studies of Broadcasting 9:67–112.
——— (1974)
"The spiral of silence: A theory of public opinion." Journal of Communication 24:43–51.
——— (1976)
"The dual climate of opinion: The influence of television in the 1976 West German Federal elections." Pp. 137–169 in M. Kaase and K. von Beyme (eds.), Elections and Parties. Beverly Hills: Sage.
——— (1984)
The Spiral of Silence. Chicago: University of Chicago Press.
Paletz, D. L., J. Y. Short, H. Baker, B. Cookman Campbell, R. J. Cooper, and R. M. Oeslander (1980)
"Polls in the media: Content, credibility and consequences." Public Opinion Quarterly 44:495–513.

Rokeach, M. (1968)
"The role of values in public opinion research." Public Opinion Quarterly 3[illegible]547–59.

Roshwalb, I., and L. Resnicoff (1971)
"The impact of endorsements and published polls on the 1970 New York Senatorial elections." Public Opinion Quarterly 35:410–14.

Shamir, J. (1986)
"Preelection polls in Israel: Structural constraints on accuracy." Public Opinion Quarterly 50:62–75.

Stone, G. C., and J. Morrison (1976)
"Content analysis as a key to purpose of community newspapers." Journalism Quarterly 53:488–94.

Stovall, J. G., and J. H. Solomon (1983)
"The poll as a news event in the 1980 presidential campaign." Public Opinion Quarterly 48:615–23.

Weimann, G. (1983a)
"Pre-election polls in the Israeli press." Journalism Quarterly 50:315–26.

——— (1983b)
"Every day is election day: Press coverage of pre-election polls in Israel." In E. Guttman, A. Diskin, and D. Caspi (eds.), The Roots of Begin's Success: The 1981 Israeli Elections. London: Croom Helm.

——— (1989)
"The obsession to forecast." Paper presented at the Esomar Conference on Public Opinion, Seville, Spain.

Wheeler, M. (1976)
Lies, Damn Lies, and Statistics: The Manipulation of Public Opinion. New York: Liveright.

Wilhoit, G., and T. Sup Auh (1974)
"Newspaper endorsement and coverage of public opinion polls in 1971." Journalism Quarterly 51:654–658.

THE POLLS—A REPORT
THE SEXUAL REVOLUTION?

TOM W. SMITH

"Sexual Revolution" became the metaphor of choice to describe the changes in sexual attitudes and behaviors during the 1960s. It triumphed over such other revolutionary standards as the "morals revolution" (*Newsweek,* April 6, 1964), the "moral revolution" (*New York Times,* May 7, 1965), the "sex revolution" (Glenn and Weaver, 1979; Sorokin, 1956), and the "sexplosion" (*Christian Century,* January 29, 1964).

A review of magazine coverage during the 1960s suggests that the Sexual Revolution was discovered by the mass media in 1963–64. In 1963 *Time* (March 22), *America* (April 20), and *Mademoiselle* (October), all referred to the Sexual Revolution. This attention culminated in *Time*'s cover story on January 24, 1964.

Of course it is possible to find earlier references to the Sexual Revolution. *Time* (January 24, 1964), in a reference to the 1920s, called the changes of the 1960s "The Second Sexual Revolution."[1] *Esquire* (July, 1961) found the revolution underway in 1961 and Pitirim Sorokin (1956) had the "American Sex Revolution" in full throttle by 1954.

But in terms of both when the mass media widely adopted the term "sexual revolution" and the way it is now typically applied by most popular and academic users, the Sexual Revolution started in the early 1960s.[2]

TOM W. SMITH is Director of the General Social Survey at the National Opinion Research Center, University of Chicago. A special thanks is owed to Patrick O'Malley for making unpublished results available.

---

[1] The 1920s have been typically depicted as a period of sexual permissiveness (and immorality in general) (Caplow et al., 1982; Reiss and Lee, 1988). They have popularly been labeled the Roaring Twenties and the Lawless Decade. Several scholars join *Time* in labeling changes during this period as a "Sexual Revolution" (Caplow et al., 1982; Ehrmann, 1961).

[2] The emergence of the Sexual Revolution also marked the appearance of the series of Poll Reports on sex and morality in *Public Opinion Quarterly* in 1966 and 1967 ("The Population Explosion, Birth Control, and Sex Education," [Erskine, 1966]; "Morality," [Erskine, 1966/67); "More on Morality and Sex," [Erskine, 1967a]; and "More on the Population Explosion and Birth Control" [Erskine, 1967b]).

Public Opinion Quarterly Volume 54:415–435 © 1990 by the American Association for Public Opinion Research
Published by The University of Chicago Press / 0033-362X/90/0054-03/$2.50

As commonly used, the term "sexual revolution" indicates a revolutionary uprooting of traditional sexual morality. It is associated with a plethora of attitudinal and behavioral changes: the new morality and the Playboy Philosophy, communes and cohabitation, free love and easy sex, wife swapping and swinging, coming out of the closet and living out of wedlock, x-rated movies and full-frontal foldouts. It has prompted the avant garde to celebrate the overthrow of a repressive Puritanism and traditionalists to lament the triumph of libertine hedonism. Whether commending or condemning the trend, few seem to question the enormous changes that the Sexual Revolution brought to America.

Then, according to the popular imagery, came the Thermidor of the 1980s. Challenged by the rise of the New Christian Right led by the Reverend Jerry Fallwell's Moral Majority, the Sexual Revolution was finally routed by the AIDS epidemic. *Time,* which had heralded the start of the Sexual Revolution on its cover in 1964, pronounced on another cover in 1984, "The revolution is over."

But was there a Sexual Revolution and did it end as the popular chronicles tell us? A review of survey data over the last four decades indicates that some important changes have occurred in the sexual arena, but raises some serious questions about the notion of the rise and fall of a Sexual Revolution.

There are time series on attitudes toward premarital, extramarital, and homosexual relations, pornography, sex education, birth control, and a few miscellaneous topics. On the one hand the data cover a wide range of topics and include both adult and youth samples (high school seniors and first-year college students). But neither the coverage of topics nor the coverage of time is complete enough to fully monitor changes in sexual attitudes. Data on sexual attitudes before 1969 are especially sparse. In the 1970s and 1980s attitudinal data concerning both adults and students are fairly extensive. There are, however, so few behavioral indicators that this dimension must be largely unexamined.[3]

Trends by Topics

Looking at each area of sexual attitudes in turn we find the following trends:

1. Premarital sex: Approval of premarital sex was stable before the 1960s. Approval began rising sometime in the 1960s and then leveled off in the early 1980s (Glenn and Weaver, 1979;

---

[3] A major analysis of trends in sexual behaviors is being currently carried out by the author. For a list of the key representative behavioral studies see Smith, 1990b. See also Klassen et al., 1989.

Smith, 1990a; Singh, 1980). Even acceptance of having children outside of marriage has gained widespread support since 1970. Among first-year college students approval of what might be characterized as casual sex has changed little since 1974, while among high school seniors approval of monogamous sex has grown. Among both high school seniors and first-year college students, approval of cohabitation was stable or in decline until the mid-1980s when it moved up to a new high. Among the former, opposition to having children out of wedlock rose in the late 1970s and early 1980s, but then fell to a new low in 1986–88.

2. Extramarital sex: Disapproval of extramarital sex was strong and stable from 1970 to 1987 and then showed a conservative shift in 1988 and 1989 (Greeley, forthcoming; Greeley, Michael, and Smith, 1989; Saunders and Edwards, 1984). Support for marriage itself has been high and steady among high school seniors, but belief in its utility (that marriage brings lasting happiness) has weakened.
3. Homosexuality: Disapproval of homosexuality has been high and has increased over time. In particular, disapproval of homosexual sex and opposition to legalization of homosexual relations grew in the mid-1980s, but some signs of a reversal appear in 1989. This trend was partially counterbalanced by growing support for civil liberties for homosexuals during the 1980s (Smith, 1988). Among first-year college students approval of laws prohibiting homosexual relations also peaked in 1987 with a reversal in 1988–89.
4. Pornography: Approval of sexually explicit materials shows a complex trend. Objections to various sexual presentations diminished in the early 1970s. By the mid-1970s, however, attitudes seem to have turned against pornography. Support for the ending of all laws regulating pornography declined and belief in the negative effects of pornography grew. In the 1980s there was little overall change, with some items showing small liberal movements and others similar conservative shifts. Viewing of x-rated movies shows its own unique pattern (Smith, 1987). From 1973 to 1978 viewing declined. This is probably related to the redefinition of the x-rating to refer only to pornographic movies designed for the adult theatre trade rather than encompassing major motion pictures intended for general release. Then from 1978 to 1987 viewing of x-rated movies climbed rapidly. This is believed to be related to increased home viewing on VCRs. Finally, in 1988 and 1989 there seems once again to be a fall-off in viewing.

5. Sex Education: Approval of sex education in schools has been growing since the 1960s. Perhaps reflecting that increased approval, the proportion of high school students taking sex education courses in general and sex education courses that discussed birth control in particular has increased since the mid-1970s.
6. Birth Control: Since the late 1950s public support for making birth control information available to adults has been rising. When the question was discontinued in 1983, approval had topped 90%. Similarly, support for making such information available to teenagers grew in the 1970s and 1980s (reaching an approval level of more than 85%). In the 1980s support for making birth control methods available to teenagers without parental consent did not change. High school seniors have shown a more mixed trend however. Since the mid-1970s, they have increased approval of teaching about birth control in sex education courses, raised and then lowered their agreement that birth control is immoral (for a small net shift to the moral side), and lowered and then raised their approval of the government making birth control freely available to all (for a net shift against free birth control).
7. Other Sexual Matters:
   a) From 1939 to 1959 there was no change in approving women wearing shorts,
   b) From 1971 to 1985 reports of engaging in "skinny dipping" increased, and
   c) Since 1978 approval of "More acceptance of sexual freedom" declined.

Trends by Decades

The limited data from before the 1960s indicate that sexual attitudes were undergoing little change. The short series on premarital sex, sex education, birth control, and proper streetwear for women are all either constant or have no clear trend. From the 1960s to the 1970s, on the other hand, the time series on premarital sex, pornography, sex education, and birth control information all show an increase in liberal or permissive attitudes. However, since only three trends (two on sex education and one on birth control information) start before 1969, we really do not know when the liberalization started or what areas it affected.

In the 1970s and early 1980s there are many more trends and the pattern is rather complex. Approval of premarital sex, sex education, birth control information for teenagers, and engaging in skinny dipping had liberal shifts. Little or no change, however, occurred on approval of extramarital sex and birth control information for adults. Conservative trends appeared on homosexuality, pornography, and sexual freedom.

The 1980s showed an even more complex pattern with series within the various topical areas often showing different trends and/or changes differing from year-to-year. Overall, there was some further diminution of liberal trends. Only sex education continued to gain public support. Little or no change occurred on premarital sex, several measures of pornography, and the availability of birth control devices for teenagers. Conservative trends appeared for extramarital sex, some measures of homosexuality, and welcoming more sexual freedom.

Conclusion

The metaphor of a Sexual Revolution captured the imagination of a generation of Americans, but poorly describes and generally exaggerates the changes in the sexual mores of Americans (Klassen, Williams, and Levitt, 1989; Niemi, Mueller, and Smith, 1989). The trends in sexual attitudes that can be tracked hardly amount to a Sexual Revolution. They are both smaller and more nuanced than aptly fits a revolutionary characterization. And without a revolution we can not have a counterrevolution. Yet on a modest level, there is support for the imagery of revolution and counterrevolution. Notable increases in approval of premarital sex (including cohabitation), sex education, and birth control did occur over the last generation. However, at least since the early 1970s there appears to have been no liberal shift, and even some conservative movement, in attitudes on homosexuality, extramarital sex, and pornography. In the early 1980s, a number of liberal trends did slow down and some reversed. However, the conservative shift was usually rather modest and in a number of cases (especially among students) shifted once again toward liberalization by the late 1980s.

SOURCES

ABC/*WP:* ABC News/*Washington Post*
CBS/*NYT:* CBS News/*New York Times*
CIRP: Cooperative Institutional Research Program, University of California, Los Angeles
Gallup: Gallup Organization
Harris: Louis Harris and Associates
LAT: Los Angeles Times
MTF: Monitoring the Future, Survey Research Center, University of Michigan

NORC: National Opinion Research Center, University of Chicago
NORC-GSS: NORC's General Social Survey, University of Chicago
RAC: Response Analysis Corporation
Roper: Roper Organization
VS: Virginia Slims surveys—Harris: 1970; Roper: 1979, 1985
YSW: Yankelovich, Skelly, and White

References

America (1963)
"Sexual revolution." America, April, 20, 526.
Boroff, David (1961)
"Sex: the quiet revolution." Esquire, July, 96–99.
Caplow, Theodore, Howard M. Bahr, Bruce Chadwick, Reuben Hill, and Margaret Holmes Williamson (1982)
Middletown Families: Fifty Years of Change and Continuity. Minneapolis: University of Minnesota Press.
Ehrmann, Winston (1961)
"Changing sexual mores." In Eli Ginzberg (ed.), Values and Ideals of American Youth. New York: Columbia University Press.
Erskine, Hazel Gaudet (1966)
"The population explosion, birth control, and sex education." Public Opinion Quarterly 30:490–501.
——— (1966/67)
"Morality." Public Opinion Quarterly 30:669–680.
——— (1967a)
"More on morality and sex." Public Opinion Quarterly 31:116–128.
——— (1967b)
"More on the population explosion and birth control." Public Opinion Quarterly 31:303–313.
Fitch, Robert E. (1964)
"The sexplosion." Christian Century, January 29, 136–138.
Glenn, Norval, and Charles N. Weaver (1979)
"Attitudes toward premarital, extramarital, and homosexual relations in the U.S. in the 1970s." Journal of Sex Research 15:108–118.
Greeley, Andrew M., Robert T. Michael, and Tom W. Smith (1989)
"A most monogamous people: Americans and their sexual Partners." GSS Topical Report, no. 17. Chicago: NORC.
Klassen, Albert D., Colin J. Williams, and Eugene E. Levitt (1989)
Sex and Morality in the U.S.: An Empirical Enquiry under the Auspices of the Kinsey Institute. Middletown, CT: Wesleyan University Press.
Klassen, Albert D., Colin J. Williams, Eugene E. Levitt, Laura Rudkin-Miniot, Heather G. Miller, and Sushama Gunjal (1989)
"Trends in premarital sexual behavior." In Charles F. Turner, Heather G. Miller, and Lincoln E. Moses (eds.), AIDS: Sexual Behavior and Intravenous Drug Use. Washington, D.C.: National Academy Press, 1989.
Leo, John (1984)
"The revolution is over." Time, April 14, 74–83.
Mademoiselle (1963)
"David Susskind's banned program: Sexual revolution in America." Mademoiselle, October, 112ff.
McClosky, Herbert and Alida Brill (1983)
Dimensions of Tolerance: What Americans Believe about Civil Liberties. New York: Sage.
Newsweek (1964)
"The Morals Revolution on the U.S. Campus." Newsweek, April 6, 52–56.

Niemi, Richard G., John Mueller, and Tom W. Smith (1989)
Trends in Public Opinion: A Compendium of Survey Data. Westport, CT: Greenwood Press.
Reiss, Ira L., and Gary R. Lee (1988)
Family Systems in America. New York: Holt, Rinehart and Winston.
Saunders, Janice Miller, and John N. Edwards (1984)
"Extramarital sexuality: A predictive model of permissive attitudes." Journal of Marriage and the Family 46:825–835.
Singh, B. K. (1980)
"Trends in attitudes to premarital sexual relations." Journal of Marriage and the Family 42:387–395.
Smith, Tom W. (1987)
"The use of public opinion data by the Attorney General's Commission on Pornography." Public Opinion Quarterly 51:249–267.
——— (1988)
"AIDS and approval of homosexuals." GSSNews No. 2, 2.
——— (1990a)
"Reiss's trend." Unpublished paper, NORC.
——— (1990b)
"Adult sexual behavior in 1989: Number of partners, frequency, and risk." Paper presented to the American Association for the Advancement of Science, New Orleans.
Sorokin, Pitirim A. (1956)
The American Sex Revolution. Boston: Porter Sargent.
Time (1963)
"Great Britain: This happy breeding." Time, March 22, 24.
Time (1964)
"Morals: The second sexual revolution." Time, January 24, 54–59.

Premarital Sexual Relations

Roper: Do you think it is all right for either or both parties to a marriage to have had previous sexual experience?

	1937	*1959*
All right	22%	22%
All right for men only	7	8
Not all right	55	54
Don't know	14	10
Refused	2	6
N	NA	NA

Note: NA = not available.

Roper: Do you consider it all right, unfortunate, or wicked when young men have sex relations before marriage?

	Women	
	1939	*1943*
All right	9.7%	5.3%
Unfortunate	35.4	43.2
Wicked	47.3	45.9
Don't know	7.6	5.6
N	NA	2,632

NORC: Please tell me whether you agree or disagree with this statement. No decent man can respect a woman who has had sexual relations before marriage.

	1950	*1953*
Agree	29.4%	33.0%
N	1,135	1,291

Gallup: There's a lot of discussion about the way morals and sex are changing in this country, here is a question that is often discussed in women's magazines. What is your view on this—do you think it is wrong for a man and woman to have sex relations before marriage, or not?

	1969	*1973*[a]
Wrong	68.8%	47%
Not wrong	21.4	43
No opinion	9.7	9
N	1,489	1,544

[a] *Gallup:* There's a lot of discussion about the way morals and sex are changing in this country. Here is a question that is often discussed in women's magazines. What are your views on this—do you think it is wrong for people to have sex relations before marriage, or not?

Gallup: There's a lot of discussion about the way morals and sexual attitudes are changing in this country. What is your opinion about this: do you think it is wrong for a man and woman to have sex relations before marriage, or not?

	1985	*1985*[a]	*1987*
Wrong	39%	36%	46%
Not wrong	52	61	48
No opinion	9	3	6
N	1,525	1,003	1,607

[a] *Roper:* A lot of people say sexual attitudes are changing in the U.S. What is your opinion about these issues—do you think it is wrong for a man and a woman to have sexual relations before marriage or not?

VS: Now, turning to another subject, let me read you some statements. For each, please tell me whether you tend to agree or disagree.

A. Premarital sex is immoral.

	1970	*1985*
Agree	59.5%	45.0%
N	3,984	4,000

B. It should be legal for adults to have children without getting married.

	1970	*1979*[a]	*1985*
Agree	12%	33%	42%
N	3,984	3,944	4,000

[a] *VS:* Let me read you some statements. It should be legal for adults to have children without getting married. Please tell me whether you tend to agree or disagree.

NORC-GSS: There's been a lot of discussion about the way morals and attitudes about sex are changing in this country. If a man and woman have sex relations before marriage, do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

	1972	*1972*[a]	*1974*	*1975*	*1977*	*1978*	*1982*	*1983*	*1985*	*1986*	*1988*	*1988*[b]	*1989*
Always wrong	35.1%	36.5%	32.0%	29.7%	30.2%	28.6%	27.7%	27.0%	27.6%	27.0%	25.6%	33%	27.1%
Almost always	11.4	10.0	12.3	11.9	9.3	11.5	8.9	9.7	8.0	8.7	10.4	12	8.4
Sometimes	23.3	27.1	22.8	23.0	22.4	19.9	20.6	23.7	19.2	22.4	21.8	26	22.1
Not wrong at all	26.2	21.8	29.7	31.5	35.6	37.8	39.8	37.8	42.0	39.2	39.5	23	39.3
Don't know	4.1	4.6	3.2	3.9	2.6	2.2	3.1	1.9	3.1	2.7	2.7	6	3.1
N	1,602	1,460	1,477	1,485	1,520	1,528	1,501	1,592	1,530	1,464	982	2,556	1,002

[a] NORC only.

[b] *Gallup:* What is your opinion about a man and a woman having sexual relations before marriage—do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

NORC-GSS: What if they are in their early teens, say 14 to 16 years old? In that case, do you think sex relations before marriage are always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

	1986	*1988*	*1989*
Always wrong	65.8%	68.6%	70.0%
Almost always	18.7	16.1	15.8
Sometimes	10.8	11.0	8.8
Not wrong at all	3.2	3.3	4.3
Don't know	1.4	1.1	1.2
N	1,466	984	1,003

Extramarital Sexual Relations

NORC-GSS: What is your opinion about a married person having sexual relations with someone other than the marriage partner—is it always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

	1970[a]	*1973*	*1974*	*1976*	*1977*	*1980*[b]	*1980*	*1982*	*1984*	*1985*	*1987*	*1988*	*1989*
Always wrong	72.3%	69.1%	73.0%	67.8%	72.4%	65.3%	69.4%	72.4%	69.8%	74.0%	73.2%	78.4%	77.3%
Almost always	14.3	14.7	11.7	15.4	13.5	15.6	15.7	13.2	18.0	13.5	15.8	12.8	12.7
Sometimes	10.8	11.5	11.4	11.3	10.0	10.6	9.7	10.3	8.8	8.5	7.8	5.5	7.2
Not wrong	2.2	4.1	2.4	4.2	3.2	3.8	3.6	2.7	2.3	1.2	1.0	1.2	1.2
Don't know	0.4	0.6	1.5	1.3	0.9	4.7	1.6	1.4	1.1	1.2	1.0	1.2	1.2
N	3,016	1,500	1,482	1,494	1,523	1,523	1,467	1,499	1,465	1,530	1,458	975	1,031

[a] NORC: And now, would you look at the card again, and tell me which answer applies if a married person has sexual intercourse with someone other than the marriage partner?

[b] Gallup.

Homosexuality

NORC-GSS: What about sexual relations between two adults of the same sex—do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

	1973	*1974*	*1976*	*1977*	*1980*	*1982*
Always wrong	70.3%	67.0%	67.1%	68.6%	69.9%	70.3%
Almost always	6.3	4.8	5.9	5.5	5.7	5.1
Sometimes	7.3	7.5	7.5	7.2	5.8	6.3
Not wrong	10.6	12.3	15.3	14.2	13.9	14.1
Other	2.1	3.4	—	—	—	—
Don't know	3.3	4.9	4.2	4.5	4.6	4.1
N	1,497	1,484	1,488	1,522	1,465	1,497

	1984	*1985*	*1987*	*1988*	*1989*
Always wrong	73.0%	74.8%	74.8%	74.0%	70.7%
Almost always	4.8	3.9	4.1	4.5	3.9
Sometimes	7.2	6.8	6.6	5.4	5.7
Not wrong	13.8	11.9	11.9	12.3	15.0
Other	—	—	—	—	—
Don't know	3.7	3.1	2.6	3.7	4.8
N	1,466	1,531	1,450	973	1,029

Gallup: Do you think homosexual relations between consenting adults should or should not be legal?

	1977	*1981*[a]	*1982*	*1985*	*1986*	*1986*	*1987*	*1989*
Legal	43%	39%	45%	44%	32%	33%	33%	47%
Not legal	43	50	39	47	57	54	55	36
No opinion	14	11	16	9	11	13	11	17
N	1,513	1,533	1,531	1,008	1,539	978	1,015	1,227

[a] ABC/*WP*.

LAT: Do you think homosexual relations between consenting adults in the privacy of their own homes should be legal or illegal?

	1986	*1987*
Legal	53%	45%
Not legal	35	41
No opinion/Refused	12	14
N	2,405	2,040

Pornography

Gallup: Do you find actors and actresses appearing in the nude in Broadway plays objectionable?

	1969	*1973*
Yes	80.5%	64.8%
No	15.8	33.6
Don't know	3.7	1.6
N	1,523	1,544

Gallup: Would you find pictures of nudes in magazines objectionable?

	1969	*1973*
Yes	72.7%	55.0%
No	24.1	43.8
Don't know	3.2	1.2
N	1,523	1,544

Gallup: Would you find topless waitresses objectionable?

	1969	*1973*
Yes	74.3%	58.7%
No	22.4	39.1
Don't know	3.3	2.3
N	1,523	1,544

NORC-GSS: The next questions are about pornography—books, movies, magazines, and photographs that show or describe sex activities. I'm going to read some opinions about the effects of looking at or reading such sexual materials. As I read each one, please tell me if you think that sexual materials do or do not have that effect.

A. Sexual materials provide information about sex.

	1970[a]	*1973*	*1975*	*1976*	*1978*	*1980*
Yes	61%	61.9%	62.2%	56.5%	61.4%	58.3%
No	27	32.6	29.2	36.4	32.6	34.4
Don't know	12	5.5	8.5	7.0	6.0	7.7
N	2,486	1,499	1,486	1,493	1,529	1,465
	1983	*1984*	*1986*	*1987*	*1988*	*1989*
Yes	59.9%	57.6%	57.0%	61.6%	58.4%	56.9%
No	35.1	37.4	37.9	32.2	35.2	38.1
Don't know	5.0	5.1	6.1	6.1	6.3	5.0
N	1,591	1,463	1,466	1,459	993	1,030

[a] *RAC* (differences in the list used and in wording makes this only partially comparable to latter points):
On this card are some opinions about the effects of looking at or reading sexual materials. As I read the letter of each one please tell me if you think sexual materials *do* or *do not* have these effects. Let's start with letter a:
a. Sexual materials provide entertainment, b. sexual materials make people bored with sexual materials, c. sexual materials provide an outlet for bottled up impulses, d. sexual materials make people sex crazy, e. sexual materials give relief to people who have sex problems, f. sexual materials lead to a breakdown of morals, g. sexual materials improve sex relations of some married couples, h. sexual materials provide information about sex, i. sexual materials excite people sexually, j. sexual materials lead people to commit rape, k. sexual materials lead people to lose respect for women, l. sexual materials make men want to do new things with their wives.

B. Sexual materials lead to breakdown of morals.

	1970[a]	*1973*	*1975*	*1976*	*1978*	*1980*
Yes	56%	52.9%	51.4%	55.0%	56.9%	60.3%
No	30	41.5	39.5	37.6	38.0	32.6
Don't know	14	5.6	9.0	7.4	5.1	7.0
N	2,486	1,500	1,487	1,490	1,529	1,461

	1983	*1984*	*1986*	*1987*	*1988*	*1989*
Yes	58.7%	61.8%	62.2%	61.9%	62.1%	62.5%
No	36.1	33.1	32.5	32.3	32.8	31.4
Don't know	5.2	5.1	5.3	5.8	5.1	6.0
N	1,589	1,462	1,463	1,459	992	1,028

[a] RAC. Differences in the list used and slight differences in wording makes this only partially comparable to latter points.

C. Sexual materials lead people to commit rape.

	1970[a]	*1973*	*1975*	*1976*	*1978*	*1980*
Yes	49%	50.0%	52.0%	52.7%	56.7%	54.0%
No	29	42.9	37.9	38.1	35.9	36.8
Don't know	22	7.0	10.0	9.1	7.3	9.2
N	2,486	1,495	1,484	1,492	1,528	1,462
	1983	*1984*	*1986*	*1987*	*1988*	*1989*
Yes	55.4%	55.1%	56.8%	54.3%	56.4%	59.9%
No	37.5	37.1	36.3	36.8	35.5	32.9
Don't know	7.1	7.8	6.9	8.9	8.1	7.2
N	1,589	1,462	1,461	1,458	989	1,029

[a] RAC. Differences in the list used and slight differences in wording makes this only partially comparable to latter points.

D. Sexual materials provide an outlet for bottled-up impulses.

	1970[a]	*1973*	*1975*	*1976*	*1978*	*1980*
Yes	34%	54.7%	56.3%	55.6%	58.9%	59.2%
No	46	35.2	27.8	28.6	29.1	26.9
Don't know	20	10.1	15.9	15.8	12.0	14.0
N	2,486	1,499	1,485	1,491	1,527	1,462
	1983	*1984*	*1986*	*1987*	*1988*	*1989*
Yes	57.4%	60.6%	59.5%	60.5%	56.1%	60.0%
No	31.5	27.3	30.6	26.6	30.5	26.3
Don't know	11.1	12.1	9.9	12.9	13.3	13.7
N	1,588	1,463	1,462	1,457	992	1,029

[a] RAC. Differences in the list used and slight differences in wording makes this only partially comparable to latter points.

NORC-GSS: Which of these statements comes closest to your feelings about pornography laws?

There should be laws against the distribution of pornography whatever the age.

There should be laws against the distribution of pornography to persons under 18.

There should be no laws forbidding the distribution of pornography.

	1973	*1975*	*1976*	*1978*	*1980*	*1983*
Laws, all ages	41.8%	40.4%	40.0%	43.2%	40.3%	41.1%
Laws, under 18	47.1	47.6	49.8	48.5	51.4	53.1
No laws	9.2	10.7	8.1	7.3	6.1	4.5
Don't know	1.9	1.2	2.1	1.0	2.2	1.3
N	1,498	1,489	1,496	1,531	1,468	1,595

	1984	*1986*	*1987*	*1988*	*1989*
Laws, all ages	40.6%	42.6%	39.6%	43.2%	40.3%
Laws, under 18	54.0	52.8	55.2	50.2	53.5
No laws	4.2	3.6	4.0	4.7	4.8
Don't know	1.2	1.0	1.2	1.9	1.4
N	1,464	1,469	1,461	994	1,032

YSW: A number of controversial proposals are being discussed these days. I'd like you to tell me for each one whether you strongly agree with it, partially agree with it, or disagree with it.

The government should crack down more on pornography in movies, books, and nightclubs.

	1974	*1975*	*1976*	*1977*	*1981*	*1982*	*1986*
Strongly agree	42%	45%	47%	56%	51%	50%	45%
Partially agree	32	19	22	20	18	22	27
Disagree	26	36	32	23	31	28	27
N	1,238	1,014	951	1,044	1,221	1,010	1,017

Roper: Here is a list of laws that have been proposed. Would you read down that list and for each tell me whether you would be for or against such a law?

A law banning pornographic book stores in residential shopping areas.

	1976	*1982*
For	70%	75%
Against	25	23
Don't know	5	3
N	2,000	2,000

Gallup: Do you think the standards in your [1977: this] community regarding the sale of sexually explicit material should be stricter than they are now, not as strict, or kept as they are now?

	1977	*1985*	*1986*
Stricter	45%	43%	45%
Less strict	6	4	5
Kept as now	35	48	43
No opinion	14	7	5
N	1,507	1,020	1,539

Roper: Here is a list of some quite different laws that could come up for a vote in any community. Would you go down the list and for each one tell me whether you would vote for it or against it if it were put to a vote in your community?

A law that would forbid the showing of any X-rated movie.

	1977	*1985*
Vote for	45%	37%
Vote against	50	61
Don't know	5	2
N	2,000	1,003

NORC-GSS: Have you seen an X-rated movie in the last year?

	1973	*1975*	*1976*	*1978*	*1980*	*1981*[a]	*1983*	*1984*	*1986*	*1987*	*1988*	*1989*
Yes	25.2%	18.8%	18.2%	14.9%	16.0%	20%	19.5%	23.5%	24.7%	27.8%	26.6%	22.7%
N	1,500	1,489	1,495	1,532	1,467	1,533	1,595	1,466	1,466	1,462	997	1,031

[a] ABC/*WP*.

Roper: Here is a list of a number of different things. Would you read down that list and call off each one you personally have done in the last month?

Been to an X-rated movie

	1978	*1985*
Yes	4%	7%
N	2,000	1,998

Sex Education

Gallup: It has been suggested that a course in sex education be given to students in high schools. Do you approve or disapprove of this?

	1943	*1951*
Agree	67.8%	66.6%
Disagree	16.9	22.6
Don't know	15.3	10.8
N	3,052	1,355

NORC-GSS: Would you be for or against sex education in the public schools?

	1965[a]	*1969*[b]	*1970*[b]	*1970*[c]	*1974*	*1975*	*1977*	*1977*[d]	*1978*[e]	*1982*	*1983*	*1985*	*1986*	*1988*	*1989*
For	68.9%	71%	65%	70%	78.9%	76.3%	77.0%	77.2%	72%	81.7%	83.6%	82.0%	81.7%	84.7%	84.5%
Against	21.9	22	28	23	17.1	20.0	20.9	16.3	23	15.2	13.6	15.5	15.5	12.6	12.2
Don't know	9.2	7	7	6	4.0	3.7	2.1	6.5	5	3.1	2.8	2.5	2.8	2.7	3.4
N	3,499	NA	NA	2,486	1,481	1,488	1,524	1,518	1,599	1,504	1,595	1,526	1,468	984	1,003

NOTE: NA = not available.

[a] Gallup: Many schools give courses in sex education. Do you approve or disapprove of such courses?

[b] Gallup: wording uncertain.

[c] RAC: By the way, would you be for or against sex education in the public schools? "For" includes qualified approval responses.

[d] Gallup: Do you approve or disapprove of schools giving courses in sex education?

[e] CBS/*NYT*: Do you approve or disapprove of schools giving courses in sex education?

Gallup: Do you approve or disapprove of schools giving courses in sex education. If approve: Would you approve or disapprove if these courses discussed birth control?

	1965[a]	*1977*	*1978*[b]
Approve	46.2%	69.4%	65.5%
Disapprove[c]	49.6	28.5	32.3
Don't know	4.3	2.1	2.2
N	3,499	1,518	1,599

[a] Gallup: Many schools give courses in sex education. Do you approve or disapprove of such courses? If approve: Would you approve or disapprove if they discussed birth control?

[b] CBS/*NYT*.

[c] Includes those who do not approve on the screening question.

Gallup: Do you feel the public elementary schools should or should not include sex education in grades 4 through 8?

	1981	*1985*	*1987*
Yes	45%	52%	55%
No	48	43	37
Don't know	7	5	8
N	1,591	1,528	1,571

Gallup: Do you feel (that/the) public high schools should or should not include sex education in their instructional program?

	1981	*1985*	*1987*
Yes	70%	75%	76%
No	22	19	16
Don't know	8	6	8
N	1,591	1,528	1,571

Birth Control

Gallup: Would you approve or disapprove of having government health clinics furnish birth control information to married people who want it?

	1940	*1943*	*1945*[a]	*1947*[a]
Approve	68.4%	61.3%	61.8%	67.4%
N	1,510	3,015	3,060	1,453

[a] Gallup: adds "in this country."

Gallup/NORC-GSS[a]: In some places in the United States, it is not legal to supply birth control information. How do you feel about this—do you think birth control information should be available to anyone who wants it, or not?

	1959	*1961*	*1962*	*1963*	*1964*	*1965*
Yes	73.4%	75.9%	72.9%	74.8%	81.4%	74.8%
No	13.6	—	—	—	11.0	15.8
Don't know	13.0	—	—	—	7.6	9.4
N	1,549	1,606	1,499	1,624	3,528	551

	1974	*1975*	*1977*	*1982*	*1983*
Yes	91.2%	89.4%	91.1%	90.4%	90.2%
No	7.7	8.5	7.3	7.5	8.0
Don't know	1.0	2.1	1.6	2.1	1.8
N	1,484	1,488	1,525	1,506	1,597

[a] Gallup prior to 1974; NORC-GSS 1974 and thereafter.

NORC-GSS: Do you think birth control information should be available to teenagers who want it, or not?

	1974	*1975*	*1977*	*1982*	*1983*
Yes	77.8%	78.7%	82.0%	86.0%	85.1%
No	19.4	18.7	16.0	11.7	12.5
Don't know	2.7	2.6	2.0	2.4	2.4
N	1,471	1,483	1,523	1,502	1,592

NORC-GSS: Do you strongly agree, agree, disagree, or strongly disagree that methods of birth control should be available to teenagers between the ages of 14 and 16 if their parents do not approve?

	1986	*1988*	*1989*
Strongly agree	22.3%	27.0%	26.9%
Agree	34.2	29.1	26.5
Disagree	24.6	22.3	24.2
Strongly disagree	16.2	17.1	19.6
Don't know	2.6	4.6	2.8
N	1,465	987	1,004

Other Sexual Matters

Gallup: Do you think it is alright for women to wear shorts on the streets?

	1939	*1951*[a]	*1955*[b]	*1961*[c]
Approve	27.1%	20.5%	27.8%	27.5%
N	1,519	1,985	1,458	3,714

[a] Gallup: Speaking in terms of their day-to-day activities, do you approve or disapprove of women in this community wearing shorts on the street?

[b] Gallup: Do you approve or disapprove of women wearing shorts in public?

[c] Gallup: Do you approve or disapprove of women wearing shorts on the streets?

Roper: Now here is a list of a number of things. Would you read down the list and for each one tell me whether you have ever done it or not? You can just read the number and answer "Yes" or "No."

Gone swimming nude or "skinny dipping" with a mixed group

	1971	*1976*	*1985*
Yes	12%	13%	18%
N	1,499	2,002	1,997

Gallup: Here are some social changes which might occur in coming years. Would you welcome these or not welcome them?

More acceptance of sexual freedom.

	1978	*1981*	*1988*
Welcome	29%	25%	22%
Not welcome	62	67	68
Don't know	9	8	10
N	1,523	1,483	2,556

Sex and High School Students

MTF: Do you think that you would prefer having a mate for most of your life, or would you not prefer having a mate?

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Definitely prefer	53.0	58.6	57.4	59.9	60.3	59.7	60.4	60.4	59.3	58.8	60.2	57.5	58.6	63.1
Probably prefer	29.4	26.5	27.7	26.8	26.8	26.7	27.5	26.5	26.7	26.6	24.7	27.7	26.3	24.2
Not sure	13.2	11.4	11.5	9.8	9.9	10.4	9.3	10.3	10.7	10.8	11.4	10.5	10.7	9.4
Probably not	3.4	2.6	2.5	2.7	2.5	2.4	2.1	2.2	2.6	2.7	2.6	3.3	3.1	2.4
Definitely not	1.0	0.9	0.9	0.8	0.5	0.8	0.7	0.8	0.7	1.0	1.0	0.9	1.3	0.9
N[a]	2,937	2,989	3,180	3,759	3,331	3,308	3,637	3,652	3,424	3,300	3,308	3,159	3,343	3,335

[a] The reported *N*s are weighted cases. Generally the raw number of cases equalled or exceeded the weighted *N*s reported here.

MTF: Most people will have fuller and happier lives if they choose legal marriage rather than staying single or just living with someone.

	1976	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	23.9	24.4	23.7	20.7	21.4	21.6	21.2	20.9	21.9	23.9	23.3	22.6	22.1
Mostly disagree	14.2	14.5	14.8	14.6	14.4	13.3	14.9	15.1	14.3	15.1	14.5	14.4	14.6
Neither	23.7	24.5	23.5	24.5	24.5	24.9	26.3	26.5	28.1	26.4	29.9	29.2	30.5
Mostly agree	16.7	13.9	16.1	17.1	16.1	15.9	16.1	16.0	15.2	15.5	15.3	15.4	15.9
Agree	21.5	22.7	21.9	23.1	23.7	24.3	21.5	21.5	20.5	19.1	17.1	18.4	16.8
N	2,959	3,081	3,738	3,283	3,251	3,620	3,589	3,321	3,268	3,244	3,048	3,343	3,290

MTF: High schools should offer instruction in birth control methods.

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
Disagree	5.9	7.7	6.6	5.9	6.4	6.4	6.4
Mostly disagree	4.5	5.2	5.0	5.3	5.0	4.5	4.1
Neither	14.8	12.8	11.5	12.7	12.6	11.4	10.4
Mostly agree	26.8	26.5	26.0	27.3	26.2	27.1	26.5
Agree	48.1	47.7	50.8	48.7	49.7	50.6	52.4
N	2,826	2,652	2,640	3,136	2,798	2,830	3,140

	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	6.0	6.7	8.0	6.4	5.7	4.4	4.9
Mostly disagree	4.3	4.2	5.2	4.6	3.5	4.3	3.7
Neither	13.6	13.8	14.6	15.7	12.6	12.7	13.4
Mostly agree	24.2	24.0	23.3	24.5	24.5	24.4	24.7
Agree	51.9	51.3	48.8	48.9	53.6	54.2	53.3
N	3,180	2,881	2,780	2,848	2,789	2,933	2,955

MTF: I personally consider most methods of birth control to be immoral.

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
Disagree	46.0	45.3	45.1	42.3	44.1	42.5	43.5
Mostly disagree	18.1	19.1	19.1	20.0	19.6	18.4	19.7
Neither	21.5	21.0	20.8	20.9	22.6	23.2	20.5
Mostly agree	6.6	7.8	7.4	8.9	6.9	7.7	8.2
Agree	7.9	6.8	7.5	7.9	6.8	8.3	8.0
N	2,795	2,613	2,619	3,143	2,766	2,803	3,116

	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	42.3	44.2	44.1	45.7	47.1	49.9	49.5
Mostly disagree	19.1	19.2	18.8	18.4	17.4	17.7	17.1
Neither	23.4	22.4	22.9	22.2	23.3	21.5	22.1
Mostly agree	7.1	7.2	7.1	6.8	6.5	5.5	5.9
Agree	8.2	7.2	7.1	7.0	5.7	5.5	5.4
N	3,137	2,861	2,775	2,831	2,778	2,917	2,943

MTF: The government should make birth control information and services available without cost to anyone who wants them.

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
Disagree	5.3	5.4	6.4	6.7	6.1	7.4	7.2
Mostly disagree	4.1	5.3	6.0	7.2	6.5	7.0	7.2
Neither	13.3	12.7	11.8	13.2	14.1	12.5	14.0
Mostly agree	24.2	24.0	24.1	25.5	23.6	26.0	24.9
Agree	53.1	52.5	51.7	47.4	49.7	47.1	46.7
N	2,810	2,636	2,608	3,134	2,768	2,803	3,115

	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	7.8	7.2	7.7	7.5	6.7	4.9	5.2
Mostly disagree	7.4	7.6	7.5	7.4	5.9	7.5	5.5
Neither	15.5	15.1	16.6	17.1	16.3	16.7	16.5
Mostly agree	23.5	25.9	24.6	23.6	25.8	24.5	23.8
Agree	45.7	44.2	43.6	44.3	45.4	46.4	49.0
N	3,177	2,875	2,774	2,843	2,780	2,927	2,946

MTF: Did you have a unit on sex education when you were in high school?

	1976	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
No and I'm glad I didn't	7.7	7.8	9.0	9.3	8.6	7.4	8.1
No and I wish I had	31.7	32.4	31.2	31.2	28.0	24.6	25.7
Yes and it was very worthwhile	20.9	24.5	22.1	21.3	22.3	27.3	26.4
Yes and it was somewhat worthwhile	28.4	26.0	28.1	29.0	30.4	32.0	29.7
Yes but it was not worthwhile	11.3	9.4	9.6	9.2	10.8	8.8	10.0
N	2,651	2,601	3,100	2,739	2,782	3,104	3,135

	1983	*1984*	*1985*	*1986*	*1987*	*1988*
No and I'm glad I didn't	8.3	8.1	8.3	8.2	7.7	6.8
No and I wish I had	23.6	23.4	21.9	19.6	20.6	19.9
Yes and it was very worthwhile	28.3	27.3	25.9	26.3	27.0	27.3
Yes and it was somewhat worthwhile	30.9	31.5	34.5	34.6	33.9	34.3
Yes but it was not worthwhile	8.8	9.7	9.4	11.4	10.7	11.6
N	2,842	2,740	2,798	2,760	2,917	2,975

MTF: Did you ever study about birth control methods in high school?

	1976	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
No and I'm glad I didn't	11.0	11.7	11.9	13.2	13.6	10.7	10.7
No and I wish I had	44.0	41.1	39.4	40.7	37.1	33.0	35.3
Yes and it was very worthwhile	21.6	24.2	23.2	21.5	21.8	28.2	25.1
Yes and it was somewhat worthwhile	17.7	18.1	19.8	19.4	21.0	22.9	22.9
Yes but it was not worthwhile	5.6	4.9	5.7	5.2	6.5	6.0	6.0
N	2,619	2,577	3,064	2,712	2,749	3,085	3,135

	1983	*1984*	*1985*	*1986*	*1987*	*1988*
No and I'm glad I didn't	11.7	11.7	11.1	10.8	9.1	8.2
No and I wish I had	31.8	28.1	27.3	26.0	28.8	24.5
Yes and it was very worthwhile	27.4	28.1	25.9	26.7	26.9	29.7
Yes and it was somewhat worthwhile	23.9	25.1	29.7	30.1	28.4	30.7
Yes but it was not worthwhile	5.2	7.0	6.0	6.5	6.9	7.0
N	2,842	2,740	2,798	2,738	2,908	2,950

MTF: It is usually a good idea for a couple to live together before getting married in order to find out whether they really get along.

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
Disagree	32.6	29.1	30.2	28.9	32.0	31.3	31.0
Mostly disagree	15.4	14.9	14.6	14.3	15.0	15.1	14.0
Neither	17.6	15.5	15.9	17.1	16.2	16.3	16.2
Mostly agree	17.2	19.0	19.9	19.7	18.4	19.4	19.9
Agree	17.3	21.5	19.4	20.2	18.4	17.9	18.8
N	2,964	2,996	3,138	3,733	3,298	3,249	3,619

	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	29.6	26.9	27.9	25.5	24.6	21.3	20.6
Mostly disagree	13.8	14.7	13.9	13.0	13.7	11.8	13.3
Neither	14.8	15.5	15.4	16.0	14.7	14.3	14.8
Mostly agree	20.0	20.3	21.5	22.8	22.6	24.2	23.0
Agree	21.8	22.6	21.7	22.9	24.4	28.5	28.4
N	3,648	3,312	3,243	3,244	3,101	3,354	3,316

MTF: Having a close intimate relationship with only one partner is too restrictive for the average person.

	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
Disagree	36.2	36.1	36.5	38.3	41.2	41.4	39.9
Mostly disagree	23.9	22.2	23.1	23.6	23.6	22.2	23.0
Neither	15.5	14.0	13.7	13.9	13.9	14.0	13.6
Mostly agree	14.6	16.2	17.1	13.9	14.9	14.0	14.3
Agree	9.8	11.4	9.6	10.3	7.9	7.6	9.4
N	2,949	2,980	3,126	3,719	3,285	3,241	3,613

	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*
Disagree	40.8	40.0	41.1	40.9	41.2	44.1	47.5
Mostly disagree	22.8	24.2	23.2	24.2	23.1	22.6	21.2
Neither	12.6	13.3	14.4	13.2	14.2	13.8	12.2
Mostly agree	14.0	14.2	12.8	13.6	13.9	12.8	12.0
Agree	9.0	8.3	8.5	8.1	7.5	6.7	7.0
N	3,636	3,296	3,230	3,241	3,090	3,342	3,309

MTF: A man and a woman who live together without being married are . . .

	1976	*1977*	*1978*	*1979*	*1980*	*1982*
Experimenting with a worthwhile alternative lifestyle	19.1	20.5	20.4	19.5	21.0	20.2
Doing their own thing and not affecting anyone else	52.3	52.4	52.8	50.8	47.6	48.9
Living in a way that could be destructive to society	8.4	7.0	6.8	8.3	8.3	8.0
Violating a basic principle of human morality	11.7	13.9	14.0	15.6	16.7	15.4
None of the above	6.3	6.2	6.0	5.9	6.4	7.4
N	2,712	2,792	3,332	2,990	2,934	3,279

	1983	*1984*	*1985*	*1986*	*1987*	*1988*
Experimenting with a worthwhile alternative lifestyle	21.5	21.0	24.0	24.8	24.9	24.8
Doing their own thing and not affecting anyone else	47.2	50.7	50.1	50.8	50.4	50.2
Living in a way that could be destructive to society	7.5	6.0	6.0	4.9	5.7	4.3
Violating a basic principle of human morality	15.9	15.1	13.0	12.4	12.1	12.8
None of the above	8.0	7.2	6.8	7.1	6.9	7.9
N	3,339	2,945	2,918	3,073	3,370	3,326

MTF: A man and a woman who decide to have and raise a child out of wedlock are . . .

	1976	*1977*	*1978*	*1979*	*1980*	*1982*
Experimenting with a worthwhile alternative lifestyle	8.5	9.3	7.9	7.0	6.9	7.1
Doing their own thing and not affecting anyone else	34.0	31.6	31.3	28.0	26.6	30.3
Living in a way that could be destructive to society	21.9	21.7	24.8	25.5	24.9	21.5
Violating a basic principle of human morality	21.9	23.7	23.0	26.6	28.7	29.7
None of the above	13.7	13.7	13.1	13.0	12.9	11.5
N	2,694	2,778	3,302	2,962	2,936	3,257

	1983	*1984*	*1985*	*1986*	*1987*	*1988*
Experimenting with a worthwhile alternative lifestyle	7.2	8.4	8.5	8.7	9.1	9.5
Doing their own thing and not affecting anyone else	29.8	33.3	36.8	38.2	37.7	37.2
Living in a way that could be destructive to society	20.0	18.3	17.2	17.2	17.5	16.8
Violating a basic principle of human morality	29.4	27.4	24.0	22.8	23.5	23.4
None of the above	13.5	12.6	13.4	13.0	12.1	13.1
N	3,011	2,924	2,892	2,778	2,990	2,937

Sex and College Students

CIRP: Mark one in each row: Agree Strongly, Agree Somewhat, Disagree Somewhat, Disagree Strongly.

First-Year College Students
(% Agree Strongly + Agree Somewhat)

A. It is important to have laws prohibiting homosexual relations.

	1976	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
% Agree	47.0	48.6	46.3	47.3	48.9	48.6	47.2
	1983	*1984*	*1985*	*1986*	*1987*	*1988*	*1989*
% Agree	49.0	47.8	47.9	52.2	53.1	49.0	45.4

B. A couple should live together for some time before deciding to get married.

	1974	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*
% Agree	45.3	47.8	48.8	48.3	45.8	44.0	43.4	42.7
	1982	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*	*1989*
% Agree	42.8	44.8	45.1	47.4	51.1	52.1	51.3	50.7

C. If two people really like each other, it's all right for them to have sex even if they've known each other only a very short time.

	1974	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*
% Agree	46.0	50.1	49.4	50.4	48.6	49.3	47.9
	1981	*1982*	*1983*	*1984*	*1987*	*1988*	*1989*
% Agree	47.0	48.2	48.8	46.8	51.9	50.0	49.7

NOTE: Based on large samples, in 1985 *N* = 192,453